JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau

Perspective

G-Quadruplexes in Neurobiology and Virology: Functional Roles and Potential Therapeutic Approaches

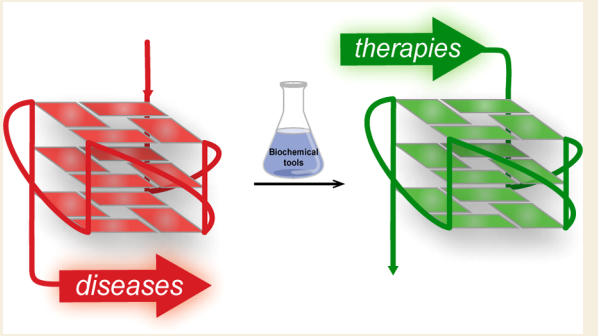
Jinglei Xu,[‡] Haiyan Huang,[‡] and Xiang Zhou*

Cite This: *JACS Au* 2021, 1, 2146–2161

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: A G-quadruplex (G4) is a four-stranded nucleic acid secondary structure maintained by Hoogsteen hydrogen bonds established between four guanines. Experimental studies and bioinformatics predictions support the hypothesis that these structures are involved in different cellular functions associated with both DNA and RNA processes. An increasing number of diseases have been shown to be associated with abnormal G4 regulation. Here, we describe the existence of G4 and then discuss G4-related pathogenic mechanisms in neurodegenerative diseases and the viral life cycle. Furthermore, we focus on the role of G4s in the design of antiviral therapy and neuropharmacology, including G4 ligands, G4-based aptamers, G4-related proteins, and CRISPR-based sequence editing, along with a discussion of limitations and insights into the prospects of this unusual nucleic acid secondary structure in therapeutics. Finally, we highlight progress and challenges in this field and the potential G4-related research fields.



KEYWORDS: *G-quadruplex, Neurology, Virology, Small molecules, RNA binding protein, Therapeutics*

INTRODUCTION

G-quadruplexes (G4s) are noncanonical guanine-rich secondary structures that exist in both DNA and RNA. The four guanine nucleotides bound through Hoogsteen hydrogen base-pairing to form a planar tetrad, called G-quartet, and two or more tetrads then interact through $\pi-\pi$ stacking to form a G-quadruplex (Figure 1A). G4s are stabilized by monovalent or divalent cations, typically potassium ions, under physiological conditions.[1] The orientation of strands and isomerization of glycosidic bonds lead to the polymorphism of G4s with parallel, antiparallel, and hybrid topological structures. The most fundamental differences between RNA and DNA G4s are the presence of uracil instead of thymine and a ribose sugar instead of a deoxyribose sugar. Combining with the presence of 2′-hydroxyl group in the ribose sugar inducing steric hindrance, resulted in anticonformation of glycosidic bond. The presence of 2′-hydroxyl group also enhanced stability for RNA G4.[2,3] Consequently, nearly all naturally occurring RNA G4s adopt a parallel topology, except for a fluorogenic RNA, Spinach, contains a G4 structure with antiparallel folding topology.[4] While DNA G4s adapt parallel, antiparallel, or hybrid conformations (Figure 1B). The distinction between DNA and RNA G4s not only displays in genome localization but also shows in regulation processing. Detailed descriptions of G4 structure and conformation are provided in other reviews.[5]

Genome-wide mapping of G4s has been achieved by computational prediction and experimental tools. On the basis of the consensus sequence features of canonical G4s, "G_3-$N_{1-7}G_3$-$N_{1-7}G_3$-$N_{1-7}G_3$ (where N = A, C, G, T or U)", researchers have developed many algorithms for computational prediction of G4s. These in silico analyses include QuadParser,[6] QGRS Mapper,[7] G4P Calculator,[8] QuadBase,[9] cGcC score,[10] and G4Hunter.[11] Considering the lack of predictions for other noncanonical G4 structures (long loops, bugles, 2 quartets), more recent computational methods added flanking sequence variants[12] and used a machine learning program[13] for better analysis. The results revealed that G4s are enriched in promoters, CpG islands, 5-untranslated regions (5′-UTRs), and nuclease-hypersensitive sites,[6,8,11] suggesting a regulatory role for G4s in key biological processes.

Experimental tools for G4 identification and analyzing their distribution in a high-throughput manner are implemented with three typical approaches: reverse transcription stalling, chemical reactive sequencing, and antibody immunoprecipita-

Received: October 12, 2021
Published: November 22, 2021

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146–2161

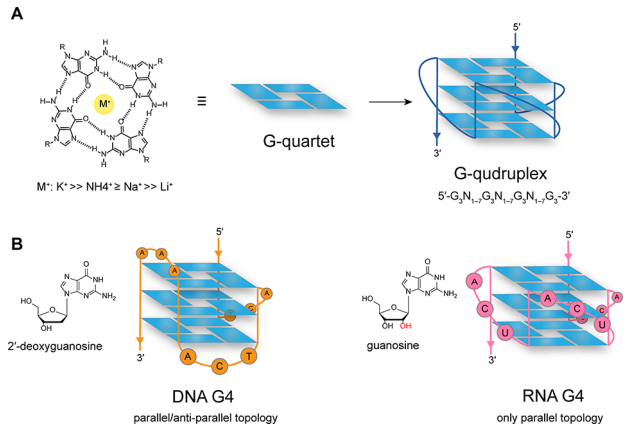


Figure 1. The G-quadruplex structure. (A) G-quartet, a planar array of four guanines stabilized by pairwise hydrogen-bonding and coordination by a cation at the center of the tetrad (M+). Preferential binding of monovalent cations to G4s and consensus sequence of G4 are shown below. (B) Differences between DNA and RNA G4 structure. Arrows indicate strand polarity. The composition of the G4 structure is telomere sequence stabilized by K^+. The 2′-hydroxyl group is highlighted in red for clarity.

tion. In the presence of potassium ions and/or G4-stabilizing ligands, G4s in DNA or RNA can cause DNA polymerase (G4-seq)[14] or reverse transcriptase stalling (rG4-seq),[15] which is combined with next-generation sequencing to enable the detection of G4s. G4-seq identified more than 700 000 G4s in DNA in the human genome, and rG4-seq showed that G4s were distributed in more than 3000 human mRNAs. G4-specific antibodies have been utilized in chromatin immunoprecipitation (ChIP) and immunohistochemical analyses to detect G4s in endogenous chromatin, and ChIP-seq peak analysis identified a positive and dynamic relationship between G4s and elevated transcriptional activity.[16] This following study identified a lower number of G4 structures (~10 000 G4 ChIP−seq peaks) than the previous G4-seq result, reflecting a suppressive role of heterochromatin for G4 formation.[17] Two chemical probes that interact with unprotected residues of RNA G4s have been applied to investigate the folding state of G4s in the cellular environment. One is dimethyl sulfate sequencing (DMS-seq). Surprisingly, DMS-seq analysis of mouse, human, and yeast cells suggests that most RNA G4 regions are unfolded.[18] The other method is selective 2′−OH acylation analyzed by primer extension (SHAPE). The SHAPE reagent reacts with all four nucleotides, which provides structural information on RNA at single-nucleotide resolution. The combination of SHAPE with lithium ions (SHALiPE) can probe RNA G4s in vitro.[19] Structure profiling of plant cells using SHALiPE showed that RNA G4s are mainly folded in vivo.[20] However, due to the complex cellular context, dynamic structural conversion, different transcript levels and heterogeneity between cells, improved methods must be developed to further confirm the folding state of RNA G4s in vivo.[21] Nevertheless, cellular G4 structures have been detected using NMR spectroscopy,[22,23] various fluorescent probes based on G4 ligands and aptamers. Other reviews provide extensive detailed description of G4 biosensing and bioimaging.[24−26] Taken together, the abundance and functional distribution of G4s, revealed using these methods, suggest that G4 motifs may confer special regulatory properties to cells.

In this Perspective, we summarize and assess the recently reported emerging roles of G4s in pathological processes of neurobiology and virology. An understanding of these roles not only provides critical insights into fundamental molecular mechanisms that control nervous system function and the viral life cycle but also provides opportunities to identify novel therapeutic targets to treat these diseases. We also discuss the recent progress in G4-targeting techniques for biomedical applications, with an emphasis on the emerging roles of G4s for examination in future studies.

FUNCTIONS OF THE G-QUADRUPLEX IN VIVO

Computational analyses of various organisms allows the prediction of putative G-quadruplex sequences (PQS), and it has revealed that the localization of PQS is nonrandom: G4 motifs usually colocalize with functional regions of the genome and are highly conserved among different species.[27] Given these results, G4 motifs may have positive functions in cells. The in vivo functions of G4s include biological processes such as telomere maintenance, DNA replication, transcriptional regulation, alternative splicing, and translational regulation. Many relatively recent studies have reported that G4 plays a functional role in phase separation,[28] epigenetic control,[29,30] and mRNA localization.[31] For a more comprehensive review on the overall topic, we refer you to other articles.[3,32,33]

The multiple functions of G4s provide an important link between G4s and many diseases, which has increased research interest in the past few years, particularly in cancer. While the biological role of G4s in cancer concentrated on their regulation of telomere maintenance, gene expression, and genome duplication, making them a potential therapeutic target against tumor cells.[34,35] However, accumulating evidence suggests important functional roles for G4 in neurological and viral diseases.[36,37] In the development and homeostasis of normal tissue, G4s have been implicated in more RNA-related mechanisms, such as mRNA splicing and localization, even regulating noncoding RNAs. Beyond different G4-related pathogenetic mechanisms of these diseases, the emerging role of G4s as therapeutic intervention of neurological disorder and antiviral is less explored. Here, we focus on recent progress in these fields, providing clarification of G4 alternative regulation from molecular mechanisms to disease states.

THE ROLE OF THE G-QUADRUPLEX IN NEUROLOGY

In recent years, G4s have been deemed to function as regulatory structural elements in many neurodegenerative diseases. Two primary mechanisms have been proposed for the contribution of G4s to neurological disorders. (i) Toxic transcribed RNA foci and dipeptide repeat proteins are generated and accumulated through the expansion of G-rich sequences (Figure 2B). (ii) Gene mutations affect pathological gene expression and G4 binding protein affinity, along with downstream regulation of defects (Figure 2A). Here, we summarize the established mechanisms of several typical neurodegenerative diseases, such as amyotrophic lateral sclerosis (ALS), frontotemporal dementia (FTD), fragile X syndrome (FXS), Alzheimer's disease (AD), and Parkinson's disease (PD), which have been proven to be related to G4 formation, with a discussion of the potential functional roles of

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

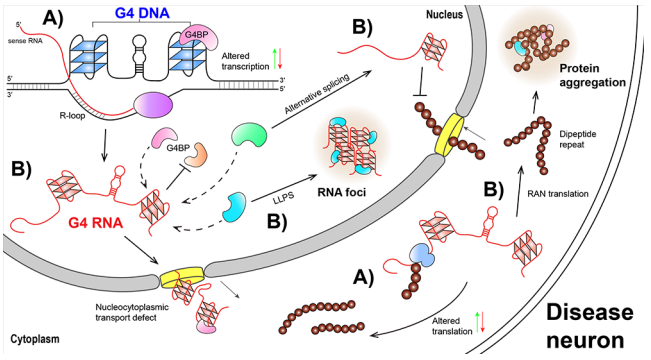


Figure 2. Proposed molecular mechanisms of G4-related neurological pathogenesis. (A) Abnormal repeat sequence in neuronal gene folded into G4 structure, inducing R-loop formation, RNA polymerase stalling, and ribosome progression stalling, resulted in genome instability, aberrant transcriptions and gene dysregulation. Multiple G4 binding protein (G4BP) involved in pathogenetic processes. (B) Transcribed RNA G4 can lead mRNA mislocalization, sequester RNA binding protein and regulate alternative splicing, cause RNA foci formation through liquid−liquid phase separation (LLPS), toxic protein aggregation through G4-induced repeat-associated non-AUG (RAN) translation, splicing dysregulation, and nucleolar stress.

G4s and future research directions. The overview of G4 in these five neurodegenerative diseases is shown in Table 1.

DNA G4 Regulates Neural Gene Transcription

The most fundamental molecular basis of neurological disorders is repeat sequence expansion in neuronal genes. For example, a familial mutation causing a G-rich sequence expansion in chromosome 9 open reading frame 72 (*C9orf72*) was identified as the prevailing genetic cause of two related neurodegenerative syndromes, FTD and ALS.[38,39] Abnormal expansion of the hexanucleotide repeat GGGGCC (G4C2) in the noncoding intron of the *C9orf72* gene is usually present in approximately 2−23 units in healthy individuals,[40] but patients with these two diseases have pathogenic G4C2 repeats ranging from 500 to 3500 units.[41,42] Several studies of G4C2 repeat DNA showed that disease-relevant DNA repeats form a mixture of parallel and antiparallel DNA G4 structures in *C9orf72* impede transcription,[43−45] inducing the generation of abortive transcripts and resulting in the loss of *C9orf72* function.[44] Researchers have shown that the G4C2 repeat also forms transcriptionally induced DNA·RNA hybrids, known as R-loops. An in vitro study showed that R-loops caused RNA polymerase stalling and the accumulation of abortive *C9orf72* transcripts. Recent studies implied that G4 motifs in the nontemplate DNA strand were induced by transcription initiation and that R-loop formation occurred prior to G4 formation and was stabilized by G4s.[46,47] The mechanisms of G4 and R-loop regulatory feedback in the *C9orf72* gene still require further investigation in the future.

In addition, DNA G4s have also been shown to the inhibit transcription of *BRCA1*, a gene encoding a critical DNA repair factor that has also been linked to Alzheimer's disease (AD). Two G4 structures in the Brca1 gene promoter are stabilized by PDS and TPMyP4, promoting neurotoxicity and the formation of DNA double−strand breaks (DSBs) in cultured neurons.[66] Recently, a ChIP-seq analysis showed that BMI1, a Polycomb group protein associated with sporadic late-onset AD, is enriched at heterochromatin regions containing putative G4 DNA sequences in neuronal cells. The accumulation of DNA G4s reduces BMI1 expression levels in AD neurons, and G4 structures essentially form only in actively transcribed LINE1s elements. This result suggests that G4 structures alter gene expression and are drastically reduced in AD neurons by more tightly compacting heterochromatin.[67]

The G4 motifs in the *C9orf72* gene also exhibit close relationships with genome stability and epigenetic regulation. G4C2 repeats lead to a decrease in replication efficiency and an increase in instability in a length-dependent manner.[48] Although G4 has been proven to cause replication fork stalling and alter epigenetic modifications, an understanding of the organizational role of G4 as a key regulator will improve therapeutic targeting in the future.

RNA G4 Regulates Neural mRNA Localization and Gene Expression

In addition to the proposed mechanisms described above that repeat expansions lead to neural gene loss of function, two other potential mechanisms of neurological disease patho-

Table 1. Overview of G-Quadruplexes in Neurological Diseases: Position, Proposed Function, and Related Diseases

gene	position	biological role of G4	disease	mechanism[a]
C9orf72	intron 1	decrease replication efficiency[48]	frontotemporal dementia/amyotrophic lateral sclerosis (FTD/ALS)	(i)
		impede transcription[44]		(i)
		influence mRNA splicing[49]		(ii)
		inducing RNA foci formation[38,44,50]		(i)
tiRNA		repress translation[51]		(ii)
FMR1	5′-UTR	repress translation,[52,53]	fragile X syndrome (FXS)	(i)
		reducing RAN translation,[54]		
		mRNA mislocalization[54,55]		
	exon 15	influence mRNA splicing[52]		(i)
PSD95	3′-UTR	repress translation[56,57]		(ii)
BACE1	exon 3	influence mRNA splicing[58,59]	Alzheimer's disease (AD)	(ii)
APP	3′-UTR	repress translation[60]		(ii)
ADAM10	5′-UTR	repress translation[61]		(ii)
APOE	exon 4	control gene expression[62]		(ii)
MIR1229		control miRNA maturation[63,64]		(i)
SNCA	5′-UTR	repress translation[65]	Parkinson's disease (PD)	(ii)
		reducing RAN translation[65]		(i)

[a]Mechanism (i) represents G4 intrinsic affect, (ii) represents G4-related regulation processes. Detailed mechanisms are summarized in the text.

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

genesis are the toxicity of nuclear RNA foci and toxic dipeptide repeat (DPR) proteins generated by repeat-associated non-AUG (RAN) translation of repeat-containing RNA. The structural role of G4s in these mechanisms is either disrupting the interaction between G4 binding proteins and disease-associated RNAs or recruiting G4 binding proteins to evaluate DPR protein levels.

Both full-length and abortive transcripts of the G_4C_2 repeat expansion fold into RNA G4s, and RNA G4 accumulation leads to an naturally liquid−liquid phase separation (LLPS) process, forming nuclear aggregates, or RNA foci, in the brains of patients with FTD/ALS.[38] These RNA G4s are specifically recognized by nucleolin (NCL). An in vivo study indicated that these G4-containing toxic nucleic acids cause NCL sequestration and mislocalization, leading to functional defects associated with nucleolar stress in patients with ALS/FTD.[44,50] In addition, several independent studies showed that G_4C_2 RNA repeats sequester several other proteins including splicing factors (hnRNP A1, hnRNP H); the mRNA export receptor ALYREF; the single strand DNA binding protein Pur α; the RNA-binding proteins hnRNPA3, hnRNP-U, hnRNP-K; and the RNA editing factor ADARB2. Extensive and detailed descriptions are provided in another review.[49]

In patients with FXS, early studies estimated that approximately 4% of brain mRNAs bind to FMRP,[68] while subsequent high-throughput methods identified several thousand mRNAs that potentially bind to FMRP.[69] Further biochemical assays only validated that G4s distinctly regulated the binding activity of FMRP to its own *Fmr1* mRNA, as well as to the *MAP1B*, *APP*, *Task3*, and *PP2Ac* mRNAs.[52,70] The G4s located in the 5′-UTR of these mRNAs are recognized and stabilized by FMRP, preventing ribosome scanning and repressing their translation. Thus, mRNAs with G4 motifs are authentic targets of FMRP. Nevertheless, FMRP might also interact indirectly with some mRNAs that do not contain G4s by either binding to adapter proteins or noncoding RNAs that recognize these mRNAs. However, some mRNAs harboring G4s in the 3′-UTR are also binding targets of FMRP, including the *Sem3F*, *NR2B*, and *Shank1* mRNAs.[53,71,72] These mRNAs have been shown to form one or more G4 structures in their 3′-UTR s that are recognized specifically by FMRP, suggesting a common mechanism of recognition. Conversely, FMRP recognizes and unwinds the G4 structure of the *Sema3F* mRNA through the RGG box domain.[71] FMRP binds to two stable intramolecular G4s within the *Shank1* mRNA and probably regulates dendritic mRNA translation and synaptic protein synthesis.[53] A very recent study further investigated the G4s in the 3′-UTR s of these transcripts by analyzing FMRP CLIP data, and the results revealed that the RGG domain of FMRP and the G4s in target mRNAs are important for efficient mRNA transport to neurites.[55]

Parkinson's disease (PD) is a common neurodegenerative disorder caused by the degeneration of neurons in the area of the brain that controls movement.[73] α-Synuclein (SNCA) accumulation has been reported to play a central role in the pathogenesis of PD. Bioinformatics analysis showed that three nonoverlapping G4 motifs are present in the 5′-UTR of *SNCA*, and mutations in two of the G4s promoted translation by either enhancing transcription or stabilizing the SNCA 5′-UTR reporter mRNA. Hence, these predicted G-q motifs, if preserved in the native *SNCA* mRNA, are likely to mediate a negative response in *SNCA* mRNA translation.[65] Taken together, these results identified G4 as an authentic target for PD therapy, but the effect of G4 regulation on PD remains to be elucidated.

The RAN translation products of FXTAS, a polyglycine-containing protein (FMRpolyG), are critical for the formation of inclusions observed in the brains of patients with FXTAS.[74] According to a recent study, FMRpolyG preferentially binds to CGG RNA G4s, which promotes the liquid-to-solid transition and aggregation of FMRpolyG, thereby eliciting neuronal dysfunction.[54] A recent study revealed that several G4 sequences in the 5′-UTRs of *PRKN* and *VPS35* mRNAs associated with Parkinson's disease were folded in vitro by performing biochemical assays. Moreover, the authors identified a new G4 binding protein, GNL1, which specifically interacts with G4s in the *VPS35* and *PRKN* mRNAs.[75]

Alzheimer's disease (AD) is the most common age-related neurodegenerative disease. The pathogenic mechanism of AD is typically characterized as a protein misfolding disease in which abnormal plaques form due to the accumulation of tau and β-amyloid (Aβ) proteins.[76] Aβ protein is produced by cleavage of the Aβ precursor protein (APP) by the β-site APP cleaving enzyme 1 (BACE1) and γ-secretase.[59] Biochemical assays revealed that a G-rich sequence within exon 3 of BACE1 folds into a G4 structure, and this G4 prevented alternative splicing by hnRNP H. Following hnRNP H knockdown, a decrease in the level of the full-length BACE1 mRNA and a decrease in Aβ production were observed.[58] Additionally, G4-related regulation of several other genes contributed to the pathogenetic ways for AD, including *APP*,[60] *ADAM10*,[61] and *APOE*.[62]

The Role of G4 in Regulating Noncoding RNAs

Sortilin-related receptor 1 (SORL1) is responsible for the processing and trafficking of Aβ. The levels of mature microRNA-1229-3p (miR-1229-3p) have been shown to regulate SORL1 translation in neurons.[77] Various biophysical techniques have shown that premature microRNA-1229 (pre-miR-1229) forms a G4 structure that coexists in equilibrium with the canonical hairpin structure, regulating the production of mature miR-1229-3p.[63] Similarly, the presence of the G4 structure in pre-miR-92b regulates the maturation of miR-92b, as G4 impedes Dicer-mediated maturation both in vivo and in vitro.[64] The *PSD-95* mRNA G-rich region folds into G4, and FMRP binds to this region that also encompasses the binding site for miR-125a. This G4 adopts multiple conformations, some increasing the accessibility of the miR-125a-binding site, suggesting that G4 regulates the access of miR-125a to its binding site.[56] A subsequent study showed that phosphorylated FMRP functions as a switch that regulates the translation of the *PSD-95* mRNA together with miR-125.[57] These findings indicate a potential role of G4-related miRNA regulatory pathways in the pathogenesis of neurological disorders.

Translation interfering tRNAs (tiRNAs) are tRNA fragments formed under stress conditions that may have roles in cancer progression. A G4 structure in tiRNA-Ala or tetramolecular G4s (formed from four individual tiRNAs) appears to mediate stress granule formation[51] and translation inhibition by interacting with Y-box binding protein 1, subsequently displacing eIF4F from mRNAs. DNA analogues of tiRNA-Ala may trigger a neuroprotective response in motor neurons, suggesting new possibilities for interventions in neurodegenerative diseases.

These studies proved that G4 represents a new avenue for neurodegeneration and brain aging research, and more G4-

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

related pathogenic mechanisms of neurological disorders might require further study.

THE ROLE OF G-QUADRUPLEX IN VIROLOGY

Viruses, which are simple acellular organisms mainly composed of nucleic acids genome and capsids, cause infectious diseases. Virus-borne infectious diseases tend to spread abruptly over wide geographical areas, such as the recent outbreak of coronavirus pneumonia 2019 (COVID-19) caused by the novel coronavirus SARS-CoV-2.[78] In addition to humans, G4 structures are also proposed to fold in viral life cycle (Figure 3). Various bioinformatics tools are now adopted to predict

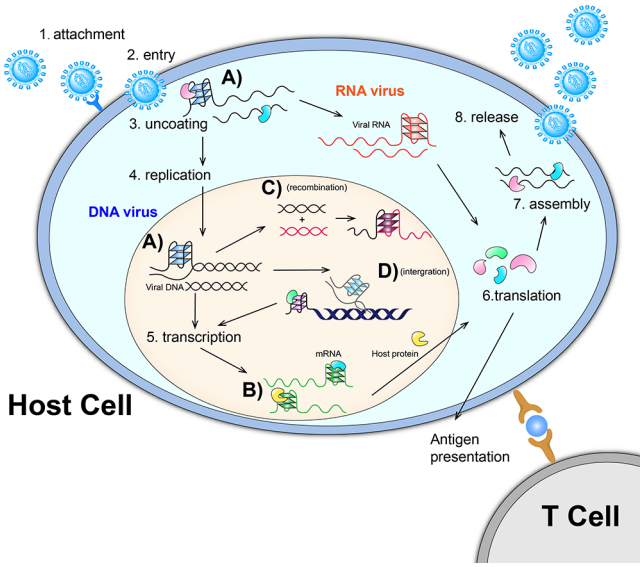


Figure 3. Role of G4 structures in viruses. Traditionally, the viral life cycle is depicted as four major steps: attachment and entry into the target cell (step 1 and 2), replication of the viral genome (step 3 and 4), translation of viral proteins (step 5 and 6) and assembly of the viral genome into infectious progeny, and egress to infect the next target cell (step 7 and 8).[81] (A) Formation of a G4 in the genome (or previral DNA) of viruses, can regulate the genome replication (or revers transcription), gene transcription, and genome integration of viruses. (B) Formation of a G4 in the mRNA of viruses can modulate the translation process, and the interaction of viral mRNA G4 with host protein or its translated protein self regulated the viral latency. (C) Intermolecular G-quartet structure increase the recombination rate of viral genome. (D) The interaction of G4s and G4-binding proteins between viruses and host, which modulate the expression of viral protein and cause viral latency.

PQS in viruses. Most of the reported viral G4s are identified first by computational prediction and then experimental tools. Two research groups have analyzed PQS in the genome of all known viruses that with the viral genome available in NCBI database.[79,80] They found that the distribution of G4s in viruses is statistically relevant, suggesting a specific biological role of G4s in viruses. The number of PQS differ across evolutionary groups and families,[79] but the distribution of PQS is highly conserved among viral strains, despite the high recombination rates of viruses.[80]

The reports that describing the presence of G4s in virus genomes have increased substantially in the recent years, and are attracting increasingly attention as a potential target for treatment and diagnosis. Several reviews have been reported focusing on the G4s in viruses.[81−84] Here, we provide an insight into the biological roles of G4s in viruses with an emphasis on the recently reported examples (all belong to human viruses), and discuss the potential of utilizing G4s as antiviral targets. The overview of G4 in viruses is summarized in Table 2.

Viral G4s Regulate Gene Replication and Expression

Under physiological conditions, the stable G4 structure leads to knot formation in the genome, obstructing replication, transcriptional (Figure 3A), and translational machinery (Figure 3B), thereby playing a major role in gene expression regulation. This pattern of regulation is reported in almost all kinds of virus.

The Herpesviridae family, which contains long linear dsDNA genomes, is extremely widespread and responsible for the onset of several diseases including different types of cancer.[85] Most of the studies conducted on the role of G4s in human herpesviruses (HHVs) focused on herpes simplex virus type 1 (HSV-1), as the genome is characterized by a 68% GC content, with multiple and highly stable G4-forming sequences.[86,87] G4 structures form in the inverted repeats region,[88] the packaging signal (pac1),[89] and the immediate early promoters,[90] have been demonstrated to may carry out regulatory functions in HSV-1. Furthermore, the HSV-1 viral protein named ICP4 was reported to bind and unfold G4s in its own promoter, thereby regulating HSV-1 viral gene expression via a G4-mediated interaction.[91] A recently reported article showed that TMPyP4, a porphyrin known to interact with G4s, displayed significant anti HSV-1 activity, which, however, was independent of inhibition of virus DNA replication or entry.[92] In addition, the stable formation of multiple clustered G4s in the repeat sequences of HSV-1 was visualized in HSV-1 infected cells using immunofluorescence and immune electronmicroscopy.[88,93]

Human papillomavirus (HPV), which is a small DNA virus with double-stranded closed loop DNA genome, led to the onset of cervical cancers and several oropharyngeal cancers. G-rich regions with the capacity to form G4s are located in the long control region (LCR) and in the coding regions of the E1, E4, and L2 proteins in eight HPV types,[94,95] all belonging to the "high risk" HPV subgroup. Therefore, G4s might play a regulatory role in these viruses.

Human immunodeficiency virus, type 1 (HIV-1) is a retrovirus that cause AIDS, which is one of the most serious public health challenges.[96] The genome of HIV-1 consists of two identical or nearly identical (+) ssRNA. Studies of G4s in the viral genomic RNA and the integrated provirus of HIV-1 have been quite productive. Several G-rich sequences were reported to form G4s in the long terminal repeats (LTRs) that enable integration of the HIV-1 provirus into the host genome.[97−100] Three overlapping quadruplex motifs were identified in the U3 region of LTRs.[101] Furthermore, G4s formed in the LTR might function as silencing elements to regulate viral transcription.[99] In addition, a bioinformation analysis of the Retroviridae family showed that most retroviruses contain highly conserved PQS in their promoter regions.[102]

The severe acute respiratory syndrome coronavirus (SARS-CoV) is a member of the Coronaviridae family and has a (+) ssRNA genome, which can lead to severe acute respiratory syndrome. Since the COVID-19 pandemic, the reports of G4s in coronaviruses have been highly productive. It has been predicted that all seven human coronaviruses harbor G-

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

Table 2. Summary of G-Quadruplexes in Viruses: Name/Genome, Position, Proposed Biological Role of G4s

virus/genome	position	biological role of G4s
HSV-1/ds DNA	inverted repeats region[88,93]	regulating viral replication and transcription[88−92]
	packaging signal (pac1)[89]	virus entry[92]
	immediate early promoters[90,91]	determine the landscape of recombination[111]
	unique long (UL) region[111]	
HPV/ds DNA	long control region (LCR)[94]	regulating viral transcription[94]
	L1, E1, and E4 coding regions[94]	
CMV/ds DNA	promoter or gene regulatory regions[112]	affecting gene expression[112]
KSHV/ds DNA	terminal repeats[113]	alter latent DNA replication and episomal persistence[113]
	LANA mRNA[114]	causing latency[113,114]
HHV-6/ds DNA	direct repeat regions[115]	modulating the interaction of viral genome into host genome to cause latency[115]
EBV/ds DNA	EBNA1 mRNA[116,117]	regulating viral replication and transcription[116−118]
		modulating immune evasion[119,120]
HIV-1/(+) ssRNA	long terminal repeats[97−101,94−98]	modulate the interaction of viral genome into host genome[97−100]
	U3 region[97,100,101]	regulating viral transcription[97,100,101]
		regulating revers transcription[101,121]
		causing latency[99,122−125]
SARS-CoV/(+) ssRNA	nucleocapsid gene[108]	repressing translation[121]
HCV/(+) ssRNA	HCV core gene[18,126]	suppressing viral gene eplication[126]
IAV/(−) ssRNA	promoter region of TMPRSS2[127]	reducing viral gene expression[127]
HBV/gapped dsDNA	envelope gene promoter[128]	regulating transcription and virion secretion[128]
	precore promoter region of cccDNA[129]	influencing HBV replication[129]

quadruplex sequences, and conserved G-quadruplex sequences in SARS-CoV and SARS-CoV-2 were analyzed and verified.[103−107] Recently, Qu et al.[108] searched for a small-molecule drug therapy by targeting the SARS-CoV-2 RNA secondary structure to fight the COVID-19 pandemic and found that PQSs located in the coding sequences region of nucleocapsid protein, named RG-1 and RG-2, are able to fold into G-quadruplexes in vitro. Furthermore, RG-1 actually forms the G4 structure in live cells, and the G4 ligand PDP stabilized RG-1 G4 and significantly reduced the level of nucleocapsid protein by inhibiting its translation both in vitro and in vivo. Thus, RG-1 G4 might be a promising therapeutic target for SARS-CoV-2.

The influenza A virus (IAV) consist of a (−) ssRNA genome, causing seasonal epidemics which can be deadly pathogens to humans. Bioinformation analysis in H1N1 influenza genomes reveals that PQS were identified in all genomes of the recently emerged genotype.[109,110] A guanine-rich tract in the promoter region of TMPRSS2 gene, which is the essential protease of IAV, was identified to form G-quadruplex in the presence of potassium ions. More importantly, compounds that are able to stabilize G4s can down-regulate TMPRSS2 gene expression as well as protein levels, and consequently suppress influenza A virus propagation in vitro.[127]

Viral G4s regulated gene replication and transcription have also been reported in the genome of human cytomegalovirus (CMV),[112] Kaposi's sarcoma associated herpesvirus (KSHV),[113] and hepatitis B virus (HBV).[128,129] In addition, PQS in the genome of HHVs is conserved in regulatory and repeated regions, corroborating that G4s are functional elements regulating the HHVs life cycle. Treatment of HSV-1-infected cells with G4 ligands such as BRACO-19 and TMPyP4, inhibit virus production.[88,92] Consequently, G4s may represent efficient pharmacological targets and alternatives to the current antiherpetic therapies.

G4s Involved in The Evolution of Viruses

The bioinformation analysis reveal that the PQS frequency, motif size, loop length, and nucleotide compositions show a correlation between the genome of virus and host, suggesting a coevolution of virus and host.[79,130,131]

HIV-1 consists of two copies of (+) ssRNA that dimerize by forming hairpin loops in the dimerization site (DIS). Several G-rich sequences have been identified in proximity to DIS, and the DIS is associated with dimerization and primer-strand transfer during reverse transcription, events that promote recombination.[121,132] Even though there's still little known about the mechanisms involved in dimerization of viral genomes, the increased rate of recombination in this region is very likely to derive from the presence of intermolecular G-quartet structure (Figure 3C).[111]

The investigation of the G4-motifs in recombination among HSV-1 found that the PQS of the genome are enriched in recombination breakpoints and in regions flanking the recombination breakpoints. The terminal guanosine of G4 clusters at the boundaries of unique long (U_L) region represent the most common breakpoint among the HSV-1 recombinants.[133] On the basis of the accumulating evidence, G4s in these regions might play a role in the evolution of viruses, suggesting that RNA G4s are fundamental for the evolution of viruses.

The Interaction of G4s and G4-Binding Proteins between Viruses and Host (Figure 3D)

Human herpesvirus 6A (HHV-6A) is a double-stranded DNA virus belong to the subfamily Betaherpesvirinae, which is the pathogen of the febrile illness roseola infantum. The HHV-6 genome is characterized by the presence of variable-length telomere-like repeat regions at its termini region, which integrate into the human chromosome at telomeres by homologous recombination. The stabilization of G4s by BRACO-19 affect the ability of HHV-6A to integrate its genome into the host chromosomes.[115] Incubation of telomerase-expressing cells with BRACO-19 caused a signifi-

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

cant reduction in the HHV-6A integration frequency; in contrast, it had no effect on the HHV-6 integration frequency in U2OS cells that lack telomerase activity. Thus, cellular G4 formation in telomeres is important for efficient chromosomal integration of HHV-6A.

Hepatitis C virus (HCV) consists of a single-stranded, positive-sense RNA genome approximately 9600 nucleotide bases in length, is a representative single-stranded RNA virus. HCV infection is a major cause of human chronic liver disease and hepatocellular carcinoma. The core gene of HCV contains a G4 RNA structure,[18] and a cellular protein, nucleolin (NCL), which was reported to bind and stabilize the HCV core RNA G4 structure.[126] HCV infection induces NCL mRNA and protein expression. While NCL suppresses wild-type viral replication, silencing of NCL substantially increases viral RNA replication. This finding provides new insights that NCL may function as a host factor for antiviral innate immunity, and binding of cellular NCL with the viral core RNA G4 structure is involved in suppressing HCV replication. Additionally, cellular nucleolin modulating EBNA1 mRNA G4s was reported, which directly mediates Epstein−Barr virus immune evasion.[119]

The SARS-unique domain (SUD), which is exclusively present in the nonstructural protein (nsp3) of SARS-CoV, is postulated to contribute to the higher pathogenicity of SARS-CoV than other human coronaviruses.[134] Interestingly, the SUD was reported to preferentially bind G4-forming oligonucleotides in the 3′-UTR of mRNAs encoding host cell proteins implicated in the regulation of several processes, such as apoptosis and signal transduction.[135] In addition, human cellular protein CNBP, which is a reported G4 helicase, was reported to bind to two PQSs in the positive-sense (+gRNA) and negative-sense (-gRNA) RNA of SARS-CoV-2.[136] Therefore, the G4 binding protein/G4 interaction between viruses and host cells is speculated to be involved in the modulation of genome replication and the host cell response to viral infection.

G4s Play a Role in Viral Latency

The latency program of viruses allows them to survive inside the host and protects them from the immune surveillance of the host. Interestingly, G4s in mRNAs of EBV and KSHV were reported to modulate the mechanism of immune evasion with two different mechanisms, suggesting that G4s may play a role in the strategies of viral latency.[114,116,117,119]

The first is that G4s in mRNA repress the translation of viral proteins, some of which exert immunomodulatory effects by restricting antigen presentation to cytotoxic T cells, allowing the virus to persist in infected cells. EBNA1 is a virus-encoded protein that is critical for the replication and maintenance of the genome during latency in proliferating cells. The EBNA1 mRNA itself is able to fold into parallel G4s.[117] The destabilization of EBNA1 mRNA G4s using antisense oligonucleotides increases EBNA1 mRNA translation. In contrast, pretreatment with a G4−stabilizing small molecule reduces EBNA1 synthesis.[116,117] Destabilization of these G4 structures limits both the presentation of MHC class I-restricted $CD8^+$ T cell epitopes by $CD11c^+$ dendritic cells in draining lymph nodes and early priming of antigen-specific $CD8^+$ T-cells.[120] In addition, cellular nucleolin, which was reported to interact with G4s in EBNA1 mRNA, was also involved in the latency of EBV.[119]

The KSHV is a human gamma herpesvirus that has been implicated in several lymphoproliferative diseases and is responsible for AIDS associated morbidities and mortalities. The genome of KSHV is organized into a 137 kb long unique region, flanked by the terminal repeats, which are rich in G residues and able to form stable G4s, both in the forward and reverse strands.[113] The latency-associated nuclear antigen 1 (LANA1) protein is the master regulatory protein of KSHV latency. It was reported that host cell protein hnRNP A1 regulate the translation of LANA mRNA through binding to the mRNA G4 structures. These findings highlight the importance of G4s within virally encoded transcripts as unique regulatory signals for translational control and immune evasion.

The second is that the viral latency related proteins with the ability to bind to G4 structures is causing viral latency. KSHV achieves latent infection by tethering its epigenome to the host chromosome by LANA, and LANA thus self-regulates its expression by mRNA sequestration in the nucleus and competing with hnRNP A1 for association with G4s at the LANA mRNA.[114] It has been reported that EBNA1 recruits the cellular origin recognition complex (ORC) through a G4 structure-related mechanism;[118] G4 interacting moleculars disrupt the interaction between EBNA1 and ORC, inhibit EBNA1-dependent stimulation of viral DNA replication, and preferentially block proliferation of EBV-positive cells.

In addition to the immune evasion, integration into the host genome is another strategy employed by HHVs for latent. As mentioned above, stabilization of telomere G-quadruplexes reduces HHV-6A chromosomal integration, suggesting that G4 structures mediate viral latency via interference with viral genome integration.[115] Similarly, G4s in HIV were also reported to modulate HIV latency with two different mechanisms: the HIV-1 LTR promoter is processed by G4 stabilizing (nucleolin)[99] and destabilizing proteins (hnRNP A2/B1),[137] and the LTR promoter has been suggested to be the region where viral latency is regulated,[122] indicating that the G4 switch may play a role in its shift to latency. Additionally, it was reported that latently infected cells incorporated silent HIV-1 provirus, resulting in dysfunctional DNA damage response (DDR). These cells are susceptible to long-term exposure to G-quadruplex stabilizing agents,[123−125] with no affect on HIV-1 promoter activity in cell culture, and this effect is enhanced when the agent is combined with an inhibitor targeting DNA-Pκ, which is crucial for repair of DNA breaks (DSB) and telomere maintenance. These findings suggest the possibility of eliminating HIV latency by targeting telomere maintenance and inhibitors targeting DDR.

THERAPIES TARGETING G-QUADRUPLEXES

As discussed above, the formation and unfolded state of G4 motifs are always in a dynamic balance, such as the potential to form G4s or block G4 structures that are not favorable for normal life processes. However, strategies that take advantage of the diverse regulatory functions of G4s by further enhancing the G4-mediated regulation of cellular processes through chemical biological tools might lead to beneficial regulation of life processes. In this section, we examine and discuss the application of G4 regulation in biomedical studies, focusing on promising techniques for targeting G4s as a therapeutic modality.

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

Targeting G4 Using Small Molecules

To date, numerous small molecules that are able to bind and stabilize G4s have been developed for G4 detection and regulation. Most of their molecular structures are similar, as planar aromatic rings interact with G-tetrads through $\pi-\pi$ stacking and electrostatic interactions. The advantage of using small molecules as therapeutics for diseases is that these compounds target not only pathogenic proteins but also cellular processes affected by the underlying pathology of G4 structures.

A set of small molecules with very similar chemical structures that target G4C2 repeat RNA G4 was found to ameliorate the two key pathologies associated with FTD/ALS in Drosophila and mammalian cell models.[138] Treatment with PDS and BRACO-19 also independently targets DNA G4 and RNA G4, resulting in the downregulation of neuronal autophagy and reducing proliferation of adult neural stem cells.[139,140] According to a previous study, oral administration of 5-ALA and TMPyP4 improves cognitive dysfunction caused by mutations in the G4-binding protein ATRX, which have been observed in a thalassemia intellectual disability X-linked syndrome mouse model and human patients. 5-ALA produces the porphyrins protoporphyrin IX (PpIX) and hemin in cells.[141] Recent studies have shown that PpIX may unfold G4s within CGG repeats in the Fmr1 transcript, resulting in the inhibition of RNA translation and ameliorating aberrant synaptic plasticity and behavior in FXTAS model mice.[54]

The stabilization of these G4s with ligands has also been investigated as a potential mechanism for inhibiting viruses (Figure 4A). Richter et al. focused on the presence and

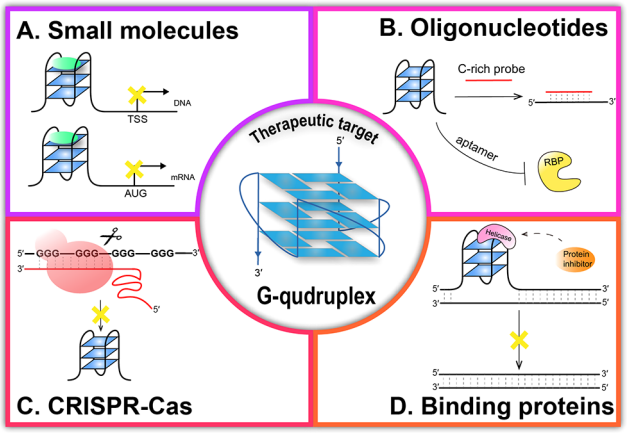


Figure 4. Potential techniques for G4-mediated therapeutic target. (A) Small molecules can specifically interact with the G4 structure, reducing transcription and translation efficiency downstream. (B) Oligonucleotide probes can form a complementary strand to G-rich sequences and prevent G4 structure formation. (C) G-rich sequences can be targeted by designed sgRNA, followed by Cas protein cleavage, abolishing G4 structure formation. (D) Regulation of G4-related protein level can influence G4 activity in cellular environment.

proposed function of G4s in virus genomes and presented the classes of G4 ligands that have been reported to display efficacious antiviral activity, with a special emphasis on the structural and physicochemical properties that characterize the viral G4/G4 ligand interaction.[81,83] For example, in HCV, TMPyP4 stabilizes RNA G4s and inhibits intracellular HCV replication at both RNA and protein levels,[18] further suppressing intracellular HCV levels in an infectious HCV culture system. Several studies have shown that small molecules that selectively target G4s in the viral genome have the potential to remove not only the replicating virus but also the latent virus, therefore preventing the development of virus-induced diseases.[117−119]

However, because of the structural similarity of RNA G4s to DNA G4s, selective targeting of RNA G4s remains challenging. To date, few selective RNA G4 ligands have been developed. Because of the conformational similarity among different G4s and the complexity of cells, complete target selectivity is difficult to achieve. In addition, small molecules have limited interaction sites, and the use of these compounds in clinical trials is far from being adopted because of their low selectivity profile and poor drug-like properties. Hence, the ability of G4 ligands to discriminate single G4 targets remains challenging in complicated cellular environments.

Targeting G4s with Oligonucleotides

Obviously, G4s are oligonucleotide-based structures, and the high specificity of base-pair recognition can be exploited as a targeting method. In contrast to quartets or tetrads, which are the "fixed" elements of the quadruplex, the nucleobase of each G4 can be quite distinct. Antisense oligonucleotides (ASOs) targeting G4 have been tested in different studies to overcome the issues with the selectivity of small molecule-based therapeutics.[142,143] Nakatani et al. reported guanine-tethered antisense oligonucleotides (g-ASs) that consisted of two functionally independent domains: one was an antisense domain that bound to the target RNA at its binding site, similar to conventional antisense DNAs, and the other was a contiguous guanine run at the 5′-end of the antisense domain. When the DNA probe hybridizes to the target, the G-rich tail composed of only three G-runs is positioned in the target to form an RNA-DNA hybrid quadruplex structure together with a guanine-rich region in the target RNA. This oligonucleotide rearrangement procedure in practice effectively inhibits reverse transcription of a variety of RNA sequences, including the HIV-1 RNA genome.[144] The remarkable ability of guanine-tethered antisense oligonucleotides to inhibit reverse transcription, translation, and replication, might facilitate the development of novel antiretroviral gene therapies based on blocking the gene replication and expression.

In contrast, an alternative approach targeting the sequence-specific disruption of a G4 was established using cytosine-rich (C-rich) nucleotide probes or and their analogues that designed to hybridize to the G-rich sequence (Figure 4B). Targeting EBNA1 mRNA with an antisense oligonucleotide complementary to the EBNA1 mRNA G4 motif destabilizes the G4 structures, stimulates EBNA1 synthesis, and enhances antigen presentation.[117] γPNA (peptide nucleic acid) oligomers, which are analogues of oligonucleotides, were reported to show high-affinity invasion to the PQS derived from the WNV NS5 gene and were previously predicted to fold into a two-tetrad RNA G-quadruplex structure.[145] After treatment of FTD/ALS cells in vitro[146,147] and in vivo with ASOs,[148] a reduction in RNA foci was observed. In addition, treatment with ASOs also led to a decrease in DPRs.[148,149]

G4-based aptamers have been identified against a wide variety of protein targets. The high affinity and selectivity for the target make aptamers very useful in many applications. Prion diseases are neurodegenerative disorders in mammals attributed to the accumulation of soluble normal cellular prion protein in the brain.[150] Researchers discovered an RNA

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

aptamer targeting bovine prion protein with a sequence of r(GGAGGAGGAGGA) (R12). It folds into a dimeric G4 structure and exhibits antiprion activity in mouse neurons by reducing the abnormal protein level.[151] Interestingly, DNA aptamers composed of two distinct classes of G4s and non-G4s forming sequences against the SARS-CoV helicase were isolated using SELEX.[152,153] While the possibility of nucleic acid therapeutics has been a promising idea for the treatment of neurological disorders, the underlying weaknesses, such as the delivery efficiency and maintenance of biostability, increase the difficulty of achieving therapeutic success.

Targeting G4 Using the CRISPR-Cas System

The CRISPR (clustered regularly interspaced short palindromic repeats)/Cas (CRISPR-associated protein) bacterial immune system has revolutionized the field of molecular biology and generated excitement for new and improved gene therapies. The simplicity and flexibility of the site-specific system has led to its widespread use in many biological research fields, including the development of model cell lines, discovery of mechanisms of disease, identification of disease targets, and transcriptional modulation.[154] As G4 possesses a consensus sequence, targeting G4 by the CRISPR-Cas system with the designed sgRNA is a reasonable approach (Figure 4C). Hence, researchers developed a CRISPR-Cas12a-based biosensor for the detection of SARS-CoV-2, as Cas12a has collateral cleavage sites in nontarget ssDNA,[155] and the authors found that the activated *Lachnospiraceae* bacterium ND2006 Cas12a (LbCas12a) also has trans-cleavage activity toward DNA G4s. Moreover, different cation concentrations resulted in efficient cleavage of G4, which was strongly related to its stability.[156] Another study investigated the structural characteristics of the complex formed by CRISPR−Cas9 and target DNA. The stability of G4 and the position of G4 in the target/nontarget strand both contributed to CRISPR−Cas9 activity.[157] As the CRISPR-Cas system has been thoroughly developed and applied in virus diagnostic tests,[158] a comprehensive approach that takes full advantage of this gene editing ability should be developed to target G4 sequences in the cellular environment. Through the elaborate design of sgRNAs with high efficiency for the target and low off-target scores, the CRISPR-Cas system will become a powerful tool for some sequence-based therapeutics for pathogenic diseases. Additionally, the addition of chemical modifications to the CRISPR-Cas system may achieve the bidirectional regulation of G4s, such as editing activated by light or chemical reagents.[159,160]

Targeting G4 Using G4-Related Proteins

The functional role of G4 is based on its efficient recognition elements for several related proteins in the cellular environment. G4 unfolding by helicase enzymes has been reported since the 1900s.[161] Interestingly, the potential of a coronavirus therapy using G4s as targets could also be achieved using molecules or aptamers that block the activity of RNA helicases (Figure 4D).[162] The nsp13 helicase in SARS-CoV unwinds the 3′- end of the nascent RNA primer, thus regulating virus replication.[163,164] The adamantane-derived bananins exhibit activity as noncompetitive inhibitors of nsp13 helicase, probably through an allosteric mechanism.[165] A novel small-molecule inhibitor was reported to block the activity of SARS coronavirus helicase by inhibiting ATP hydrolysis and dsDNA unwinding activities, without any cytotoxicity up to 80 μM.[166] In addition, researchers found that the G4 helicase DHX36 has a high affinity for G4C2 repeat RNA and efficiently unwinds the G4 structure; knockdown of DHX36 significantly decreases the levels of DPR proteins. The observation of aberrant DHX36 upregulation in tissues from patients with C9orf72-linked ALS indicates that DHX36 is a positive regulator of C9orf72 repeat-associated RNA translation.[167] Conversely, another RNA helicase, DDX3X, was found to be a negative regulator of C9orf72 repeat-associated RNA translation, and a reduction in DDX3X expression increased DPR levels in cells from patients with ALS/FTD and enhanced G4C2-mediated toxicity in *Drosophila*.[168] While DDX3X recognizes the same hairpin in the G4C2 repeat, the different effects of various secondary structures remain controversial.

As G4-binding proteins are important for regulating function, the identification of new G4-related proteins involved in these diseases will improve diagnosis and therapy. Although many proteins identified in neuronal and viral diseases tend to interact with G4s in 5′- and 3′-UTRs, a recent study identified 15 proteins whose binding sites are enriched for PQS in two cell lines, several of which have been reported to be involved in viral gene expression.[169] Further studies of these proteins will reveal the possible roles of G4s in the mechanism for coopting host cell machinery.

CONCLUSIONS AND OUTLOOK

In the past few years, G4s have attracted increasing attention, as they represent some of the most important nucleic acid secondary structures with extensive regulatory functions in cells. This widespread utilization of G4s provides an opportunity for these structures to contribute to the molecular pathology of diseases, including neurodegenerative diseases and viral pathogenesis. The common regulatory mechanisms described in this Perspective highlight several factors involved in G4s based on their structural composition, localization, and binding proteins. On the basis of the different regulatory roles of G4, we can develop different techniques for effective targeted therapies.

The flat G-tetrads of G4s enable structure-based recognition by small molecules. G4-based drugs may represent a significant turning point in the management of viral infections, especially in people who cannot access immunization, such as immunocompromised patients or elderly people. Although a number of G4 ligands have been developed, few derivatives are bioavailable and can be further utilized in clinical trials. As G4s are present in both cellular and virus genomes, one challenge is developing specific G4 ligands for pathogenic structures that do not recognize cellular G4s. Another challenge is validating the mechanism of G4-involved pathological processes. Once the effect of G4s is determined, the time points for antiviral reagent delivery and therapeutic interventions can be further investigated. Moreover, G4 is considered related to i-motif formation, epigenetic modification, and oxidative stress, which are related to the development of several diseases. Therefore, an understanding of their complex and dynamic mechanisms might provide insights into important therapeutic strategies.

Our recommendations for future research in these fields are as follows: (i) visualize the formation of G4s, (ii) explore G4 interaction network, (iii) investigate the function of G4s, and (iv) specifically manipulate G4s. First, a prerequisite is to confirm G4 structure formation in critical regions of genes associated with these diseases. Although specific antibodies, in-cell NMR analysis and many fluorescent probes are available to detect G4s in the cellular environment, contrasting observa-

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

tions of the G4 dynamic folded state have been reported in vivo,[21−23] despite different dynamic situations, cell cycles, and populations. PQS with various sequences may form different topological structures in vitro and in vivo, including parallel, antiparallel, hybrid, or interact with each other to form a higher-order structures.[170,171] Therefore, further studies utilizing newer technologies may be needed to unambiguously detect G4s in vivo and in real time. Second, DNA and RNA G4 distribution in cells have been studied by multiple methods, the formation of G4 demonstrated to be related with cellular stress,[28] molecular crowding,[172] and torsional strain,[16] but without general observation about which conditions do influence G4s formation in cells. Moreover, the key protein regulators that mediate G4 formation and depletion remain unknown. There might be interactions between G4s and other epigenetic features to form this complex regulation system. These crucial questions will help us to get a new insight for G4 biological roles. Third, based on a model in which structure determines function and function influences phenotype, DNA/RNA binding proteins either modulate the G4 conformation or serve to recruit additional protein regulators that block the processing machinery. Few efficient methods are available to identify G4-related proteins in a high-throughput manner, as most pull-down assays are performed in vitro based on G4 sequences.[173,174] We tried to perform G4-related proteomics in vivo on the basis of the G4 ligand,[175] but further study is needed to reveal the G4-protein recognition model and protein interaction network. Additionally, further study is needed to reveal the G4-protein recognition model and protein interaction network. A better understanding of the G4 binding protein will help define the regulatory role of G4 and elucidate the mechanism. Similarly, therapeutic approaches developed to manipulate the activities of *bona fide* G4-binding proteins will be promising. Fourth, manipulation of G4 formation in pathogenetic process deemed to be effective for therapeutic interventions, while the potential problems are target specificity and biocompatibility. Moreover, it is unclear whether the compounds actually affect the structure or just impair their detection. Unlike acidic extracellular pH and the limitless replication potential of tumor cells, there are fewer differences between diseased neurons and normal cells for molecule recognition. Furthermore, viral infection will trigger a robust and usually well-coordinated immune response including several pathways, which get weakened during cancer treatment. Developing a prodrug with low off-target possibilities and high delivery efficiency to the target region will meet therapeutics needed for advancement of this area.

AUTHOR INFORMATION

Corresponding Author

Xiang Zhou − *The Institute of Advanced Studies, Key Laboratory of Biomedical Polymers-Ministry of Education, Wuhan University, Wuhan 430072, China; Key Laboratory of Biomedical Polymers-Ministry of Education, College of Chemistry and Molecular Sciences, Wuhan University, Wuhan 430072, China;* orcid.org/0000-0002-1829-9368; Email: xzhou@whu.edu.cn

Authors

Jinglei Xu − *The Institute of Advanced Studies, Key Laboratory of Biomedical Polymers-Ministry of Education, Wuhan University, Wuhan 430072, China;* orcid.org/0000-0002-2240-4001

Haiyan Huang − *Key Laboratory of Biomedical Polymers-Ministry of Education, College of Chemistry and Molecular Sciences, Wuhan University, Wuhan 430072, China*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.1c00451

Author Contributions

‡J.X. and H.H. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors are grateful for financial support of the National Natural Science Foundation of China (21432008, 91753201, and 21721005 to X.Z.) and China Postdoctoral Science Foundation (2021M692468 to H.-Y.H.).

REFERENCES

(1) Guiset Miserachs, H.; Donghi, D.; Börner, R.; Johannsen, S.; Sigel, R. K. O. Distinct differences in metal ion specificity of RNA and DNA G-quadruplexes. *JBIC, J. Biol. Inorg. Chem.* **2016**, *21* (8), 975−986.

(2) Asamitsu, S.; Takeuchi, M.; Ikenoshita, S.; Imai, Y.; Kashiwagi, H.; Shioda, N. Perspectives for Applying G-Quadruplex Structures in Neurobiology and Neuropharmacology. *Int. J. Mol. Sci.* **2019**, *20* (12), 2884.

(3) Fay, M. M.; Lyons, S. M.; Ivanov, P. RNA G-Quadruplexes in Biology: Principles and Molecular Mechanisms. *J. Mol. Biol.* **2017**, *429* (14), 2127−2147.

(4) Warner, K. D.; Chen, M. C.; Song, W.; Strack, R. L.; Thorn, A.; Jaffrey, S. R.; Ferré-D'Amaré, A. R. Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Nat. Struct. Mol. Biol.* **2014**, *21* (8), 658−663.

(5) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34* (19), 5402−5415.

(6) Huppert, J. L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33* (9), 2908−2916.

(7) Kikin, O.; D'Antonio, L.; Bagga, P. S. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **2006**, *34*, W676−W682.

(8) Eddy, J.; Maizels, N. Gene function correlates with potential for G4 DNA formation in the human genome. *Nucleic Acids Res.* **2006**, *34* (14), 3887−3896.

(9) Yadav, V. K.; Abraham, J. K.; Mani, P.; Kulshrestha, R.; Chowdhury, S. QuadBase: genome-wide database of G4 DNA–occurrence and conservation in human, chimpanzee, mouse and rat promoters and 146 microbes. *Nucleic Acids Res.* **2007**, *36*, D381−D385.

(10) Beaudoin, J.-D.; Jodoin, R.; Perreault, J.-P. New scoring system to identify RNA G-quadruplex folding. *Nucleic Acids Res.* **2014**, *42* (2), 1209−1223.

(11) Bedrat, A.; Lacroix, L.; Mergny, J.-L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.* **2016**, *44* (4), 1746−1759.

(12) Garant, J.-M.; Perreault, J.-P.; Scott, M. S. Motif independent identification of potential RNA G-quadruplexes by G4RNA screener. *Bioinformatics* **2017**, *33* (22), 3532−3537.

(13) Sahakyan, A. B.; Chambers, V. S.; Marsico, G.; Santner, T.; Di Antonio, M.; Balasubramanian, S. Machine learning model for sequence-driven DNA G-quadruplex formation. *Sci. Rep.* **2017**, *7*, 14535.

(14) Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.; Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **2015**, *33* (8), 877−881.

(15) Kwok, C. K.; Marsico, G.; Sahakyan, A. B.; Chambers, V. S.; Balasubramanian, S. rG4-seq reveals widespread formation of G-quadruplex structures in the human transcriptome. *Nat. Methods* **2016**, *13* (10), 841−844.
(16) Hänsel-Hertsch, R.; Beraldi, D.; Lensing, S. V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; Tannahill, D.; Balasubramanian, S. G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.* **2016**, *48* (10), 1267−1272.
(17) Hansel-Hertsch, R.; Spiegel, J.; Marsico, G.; Tannahill, D.; Balasubramanian, S. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* **2018**, *13* (3), 551−564.
(18) Wang, S.-R.; Min, Y.-Q.; Wang, J.-Q.; Liu, C.-X.; Fu, B.-S.; Wu, F.; Wu, L.-Y.; Qiao, Z.-X.; Song, Y.-Y.; Xu, G.-H. A highly conserved G-rich consensus sequence in hepatitis C virus core gene represents a new anti−hepatitis C target. *Sci. Adv.* **2016**, *2* (4), No. e1501535.
(19) Kwok, C. K.; Sahakyan, A. B.; Balasubramanian, S. Structural Analysis using SHALiPE to Reveal RNA G-Quadruplex Formation in Human Precursor MicroRNA. *Angew. Chem., Int. Ed.* **2016**, *55* (31), 8958−8961.
(20) Yang, X.; Cheema, J.; Zhang, Y.; Deng, H.; Duncan, S.; Umar, M. I.; Zhao, J.; Liu, Q.; Cao, X.; Kwok, C. K.; Ding, Y. RNA G-quadruplex structures exist and function in vivo in plants. *Genome Biol.* **2020**, *21* (1), 226.
(21) Kwok, C. K.; Marsico, G.; Balasubramanian, S. Detecting RNA G-Quadruplexes (rG4s) in the Transcriptome. *Cold Spring Harbor Perspect. Biol.* **2018**, *10* (7), No. a032284.
(22) Salgado, G. F.; Cazenave, C.; Kerkour, A.; Mergny, J.-L. G-quadruplex DNA and ligand interaction in living cells using NMR spectroscopy. *Chem. Sci.* **2015**, *6* (6), 3314−3320.
(23) Dzatko, S.; Krafcikova, M.; Hänsel-Hertsch, R.; Fessl, T.; Fiala, R.; Loja, T.; Krafcik, D.; Mergny, J.-L.; Foldynova-Trantirkova, S.; Trantirek, L. Evaluation of the Stability of DNA i-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chem., Int. Ed.* **2018**, *57* (8), 2165−2169.
(24) Vummidi, B. R.; Alzeer, J.; Luedtke, N. W. Fluorescent Probes for G-Quadruplex Structures. *ChemBioChem* **2013**, *14* (5), 540−558.
(25) Xu, J.; Jiang, R.; He, H.; Ma, C.; Tang, Z. Recent advances on G-quadruplex for biosensing, bioimaging and cancer therapy. *TrAC, Trends Anal. Chem.* **2021**, *139*, 116257.
(26) Umar, M. I.; Ji, D.; Chan, C.-Y.; Kwok, C. K. G-Quadruplex-Based Fluorescent Turn-On Ligands and Aptamers: From Development to Applications. *Molecules* **2019**, *24* (13), 2416.
(27) Konig, S. L.; Evans, A. C.; Huppert, J. L. Seven essential questions on G-quadruplexes. *Biomol. Concepts* **2010**, *1* (2), 197−213.
(28) Zhang, Y.; Yang, M.; Duncan, S.; Yang, X.; Abdelhamid, M. A. S.; Huang, L.; Zhang, H.; Benfey, P. N.; Waller, Z. A. E.; Ding, Y. G-quadruplex structures trigger RNA phase separation. *Nucleic Acids Res.* **2019**, *47* (22), 11746−11754.
(29) Mao, S.-Q.; Ghanbarian, A. T.; Spiegel, J.; Martínez Cuesta, S.; Beraldi, D.; Di Antonio, M.; Marsico, G.; Hänsel-Hertsch, R.; Tannahill, D.; Balasubramanian, S. DNA G-quadruplex structures mold the DNA methylome. *Nat. Struct. Mol. Biol.* **2018**, *25* (10), 951−957.
(30) Fleming, A. M.; Nguyen, N. L. B.; Burrows, C. J. Colocalization of m6A and G-Quadruplex-Forming Sequences in Viral RNA (HIV, Zika, Hepatitis B, and SV40) Suggests Topological Control of Adenosine N6-Methylation. *ACS Cent. Sci.* **2019**, *5* (2), 218−228.
(31) Subramanian, M.; Rage, F.; Tabet, R.; Flatter, E.; Mandel, J.-L.; Moine, H. G−quadruplex RNA structure as a signal for neurite mRNA targeting. *EMBO Rep.* **2011**, *12* (7), 697−704.
(32) Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020**, *21* (8), 459−474.
(33) Hänsel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (5), 279−284.
(34) Balasubramanian, S.; Hurley, L. H.; Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discovery* **2011**, *10* (4), 261−275.
(35) Düchler, M. G-quadruplexes: targets and tools in anticancer drug design. *J. Drug Target.* **2012**, *20* (5), 389−400.
(36) Simone, R.; Fratta, P.; Neidle, S.; Parkinson, G. N.; Isaacs, A. M. G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Lett.* **2015**, *589* (14), 1653−1668.
(37) Wang, E.; Thombre, R.; Shah, Y.; Latanich, R.; Wang, J. G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res.* **2021**, *49* (9), 4816−4830.
(38) DeJesus-Hernandez, M.; Mackenzie, I. R.; Boeve, B. F.; Boxer, A. L.; Baker, M.; Rutherford, N. J.; Nicholson, A. M.; Finch, N. A.; Flynn, H.; Adamson, J.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G.-Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72* (2), 245−256.
(39) Renton, A. E.; Majounie, E.; Waite, A.; Simon-Sanchez, J.; Rollinson, S.; Gibbs, J. R.; Schymick, J. C.; Laaksovirta, H.; van Swieten, J. C.; Myllykangas, L.; Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J. B.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V.-M.; Kaivorinne, A.-L.; Holtta-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wuu, J.; Chio, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J. D.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.; Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72* (2), 257−268.
(40) Rutherford, N. J.; Heckman, M. G.; DeJesus-Hernandez, M.; Baker, M. C.; Soto-Ortolaza, A. I.; Rayaprolu, S.; Stewart, H.; Finger, E.; Volkening, K.; Seeley, W. W.; Hatanpaa, K. J.; Lomen-Hoerth, C.; Kertesz, A.; Bigio, E. H.; Lippa, C.; Knopman, D. S.; Kretzschmar, H. A.; Neumann, M.; Caselli, R. J.; White, C. L.; Mackenzie, I. R.; Petersen, R. C.; Strong, M. J.; Miller, B. L.; Boeve, B. F.; Uitti, R. J.; Boylan, K. B.; Wszolek, Z. K.; Graff-Radford, N. R.; Dickson, D. W.; Ross, O. A.; Rademakers, R. Length of normal alleles of C9ORF72 GGGGCC repeat do not influence disease phenotype. *Neurobiol. Aging* **2012**, *33* (12), 2950.e5.
(41) Beck, J.; Poulter, M.; Hensman, D.; Rohrer, J. D.; Mahoney, C. J.; Adamson, G.; Campbell, T.; Uphill, J.; Borg, A.; Fratta, P.; Orrell, R. W.; Malaspina, A.; Rowe, J.; Brown, J.; Hodges, J.; Sidle, K.; Polke, J. M.; Houlden, H.; Schott, J. M.; Fox, N. C.; Rossor, M. N.; Tabrizi, S. J.; Isaacs, A. M.; Hardy, J.; Warren, J. D.; Collinge, J.; Mead, S. Large C9orf72 Hexanucleotide Repeat Expansions Are Seen in Multiple Neurodegenerative Syndromes and Are More Frequent Than Expected in the UK Population. *Am. J. Hum. Genet.* **2013**, *92* (3), 345−353.
(42) Chen, Y.; Lin, Z.; Chen, X.; Cao, B.; Wei, Q.; Ou, R.; Zhao, B.; Song, W.; Wu, Y.; Shang, H.-F. Large C9orf72 repeat expansions are seen in Chinese patients with sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* **2016**, *38*, 217.e15.
(43) Šket, P.; Pohleven, J.; Kovanda, A.; Štalekar, M.; Župunski, V.; Zalar, M.; Plavec, J.; Rogelj, B. Characterization of DNA G-quadruplex species forming from C9ORF72 G4C2-expanded repeats associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **2015**, *36* (2), 1091−1096.
(44) Haeusler, A. R.; Donnelly, C. J.; Periz, G.; Simko, E. A. J.; Shaw, P. G.; Kim, M.-S.; Maragakis, N. J.; Troncoso, J. C.; Pandey, A.;

Sattler, R.; Rothstein, J. D.; Wang, J. C9orf72 nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507* (7491), 195−200.
(45) Zhou, B.; Liu, C.; Geng, Y.; Zhu, G. Topology of a G-quadruplex DNA formed by C9orf72 hexanucleotide repeats associated with ALS and FTD. *Sci. Rep.* **2015**, *5* (1), 16673.
(46) Zhao, Y.; Zhang, J.-y.; Zhang, Z.-y.; Tong, T.-j.; Hao, Y.-h.; Tan, Z. Real-Time Detection Reveals Responsive Cotranscriptional Formation of Persistent Intramolecular DNA and Intermolecular DNA:RNA Hybrid G-Quadruplexes Stabilized by R-Loop. *Anal. Chem.* **2017**, *89* (11), 6036−6042.
(47) Lee, C.-Y.; McNerney, C.; Ma, K.; Zhao, W.; Wang, A.; Myong, S. R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nat. Commun.* **2020**, *11* (1), 3392.
(48) Thys, R. G.; Wang, Y.-H. DNA Replication Dynamics of the GGGGCC Repeat of the C9orf72 Gene*. *J. Biol. Chem.* **2015**, *290* (48), 28953−28962.
(49) Vatovec, S.; Kovanda, A.; Rogelj, B. Unconventional features of C9ORF72 expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **2014**, *35* (10), 2421.e1.
(50) Wang, J.; Haeusler, A. R.; Simko, E. A. J. Emerging role of RNA●DNA hybrids in C9orf72-linked neurodegeneration. *Cell Cycle* **2015**, *14* (4), 526−532.
(51) Lyons, S. M.; Gudanis, D.; Coyne, S. M.; Gdaniec, Z.; Ivanov, P. Identification of functional tetramolecular RNA G-quadruplexes derived from transfer RNAs. *Nat. Commun.* **2017**, *8* (1), 1127.
(52) Santoro, M. R.; Bray, S. M.; Warren, S. T. Molecular Mechanisms of Fragile X Syndrome: A Twenty-Year Perspective. *Annu. Rev. Pathol.: Mech. Dis.* **2012**, *7* (1), 219−245.
(53) Zhang, Y.; Gaetano, C. M.; Williams, K. R.; Bassell, G. J.; Mihailescu, M. R. FMRP interacts with G-quadruplex structures in the 3′-UTR of its dendritic target Shank1 mRNA. *RNA Biol.* **2014**, *11* (11), 1364−1374.
(54) Asamitsu, S.; Yabuki, Y.; Ikenoshita, S.; Kawakubo, K.; Kawasaki, M.; Usuki, S.; Nakayama, Y.; Adachi, K.; Kugoh, H.; Ishii, K.; Matsuura, T.; Nanba, E.; Sugiyama, H.; Fukunaga, K.; Shioda, N. CGG repeat RNA G-quadruplexes interact with FMRpolyG to cause neuronal dysfunction in fragile X-related tremor/ataxia syndrome. *Sci. Adv.* **2021**, *7* (3), No. eabd9440.
(55) Goering, R.; Hudish, L. I.; Guzman, B. B.; Raj, N.; Bassell, G. J.; Russ, H. A.; Dominguez, D.; Taliaferro, J. M. FMRP promotes RNA localization to neuronal projections through interactions between its RGG domain and G-quadruplex RNA sequences. *eLife* **2020**, *9*, No. e52621.
(56) Stefanovic, S.; Bassell, G. J.; Mihailescu, M. R. G quadruplex RNA structures in PSD-95 mRNA: potential regulators of miR-125a seed binding site accessibility. *RNA* **2015**, *21* (1), 48−60.
(57) DeMarco, B.; Stefanovic, S.; Williams, A.; Moss, K. R.; Anderson, B. R.; Bassell, G. J.; Mihailescu, M. R. FMRP - G-quadruplex mRNA - miR-125a interactions: Implications for miR-125a mediated translation regulation of PSD-95 mRNA. *PLoS One* **2019**, *14* (5), No. e0217275.
(58) Fisette, J.-F.; Montagna, D. R.; Mihailescu, M.-R.; Wolfe, M. S. A G-Rich element forms a G-quadruplex and regulates BACE1 mRNA alternative splicing. *J. Neurochem.* **2012**, *121* (5), 763−773.
(59) Singer, O.; Marr, R. A.; Rockenstein, E.; Crews, L.; Coufal, N. G.; Gage, F. H.; Verma, I. M.; Masliah, E. Targeting BACE1 with siRNAs ameliorates Alzheimer disease neuropathology in a transgenic model. *Nat. Neurosci.* **2005**, *8* (10), 1343−1349.
(60) Crenshaw, E.; Leung, B. P.; Kwok, C. K.; Sharoni, M.; Olson, K.; Sebastian, N. P.; Ansaloni, S.; Schweitzer-Stenner, R.; Akins, M. R.; Bevilacqua, P. C.; Saunders, A. J. Amyloid Precursor Protein Translation Is Regulated by a 3′UTR Guanine Quadruplex. *PLoS One* **2015**, *10* (11), No. e0143160.
(61) Lammich, S.; Kamp, F.; Wagner, J.; Nuscher, B.; Zilow, S.; Ludwig, A.-K.; Willem, M.; Haass, C. Translational Repression of the Disintegrin and Metalloprotease ADAM10 by a Stable G-quadruplex Secondary Structure in Its 5′-Untranslated Region*. *J. Biol. Chem.* **2011**, *286* (52), 45063−45072.
(62) Chaudhary, S.; Kaushik, M.; Ahmed, S.; Kukreti, R.; Kukreti, S. Structural Switch from Hairpin to Duplex/Antiparallel G-Quadruplex at Single-Nucleotide Polymorphism (SNP) Site of Human Apolipoprotein E (APOE) Gene Coding Region. *ACS Omega* **2018**, *3* (3), 3173−3182.
(63) Imperatore, J. A.; Then, M. L.; McDougal, K. B.; Mihailescu, M. R. Characterization of a G-Quadruplex Structure in Pre-miRNA-1229 and in Its Alzheimer's Disease-Associated Variant rs2291418: Implications for miRNA-1229 Maturation. *Int. J. Mol. Sci.* **2020**, *21* (3), 767.
(64) Mirihana Arachchilage, G.; Dassanayake, A. C.; Basu, S. A Potassium Ion-Dependent RNA Structural Switch Regulates Human Pre-miRNA 92b Maturation. *Chem. Biol.* **2015**, *22* (2), 262−272.
(65) Koukouraki, P.; Doxakis, E. Constitutive translation of human α-synuclein is mediated by the 5′-untranslated region. *Open Biol.* **2016**, *6* (4), 160022−160022.
(66) Moruno-Manchon, J. F.; Koellhoffer, E. C.; Gopakumar, J.; Hambarde, S.; Kim, N.; McCullough, L. D.; Tsvetkov, A. S. The G-quadruplex DNA stabilizing drug pyridostatin promotes DNA damage and downregulates transcription of Brca1 in neurons. *Aging* **2017**, *9* (9), 1957−1970.
(67) Hanna, R.; Flamier, A.; Barabino, A.; Bernier, G. G-quadruplexes originating from evolutionary conserved L1 elements interfere with neuronal gene expression in Alzheimer's disease. *Nat. Commun.* **2021**, *12* (1), 1828.
(68) Ashley, C. T.; Wilkinson, K. D.; Reines, D.; Warren, S. T. FMR1 protein: conserved RNP family domains and selective RNA binding. *Science* **1993**, *262* (5133), 563.
(69) Chen, E.; Joseph, S. Fragile X mental retardation protein: A paradigm for translational control by RNA-binding proteins. *Biochimie* **2015**, *114*, 147−154.
(70) Maltby, C. J.; Schofield, J. P. R.; Houghton, S. D.; O'Kelly, I.; Vargas-Caballero, M.; Deinhardt, K.; Coldwell, M. J. A 5′ UTR GGN repeat controls localisation and translation of a potassium leak channel mRNA through G-quadruplex formation. *Nucleic Acids Res.* **2020**, *48* (17), 9822−9839.
(71) Menon, L.; Mihailescu, M.-R. Interactions of the G quartet forming semaphorin 3F RNA with the RGG box domain of the fragile X protein family. *Nucleic Acids Res.* **2007**, *35* (16), 5379−5392.
(72) Stefanovic, S.; DeMarco, B. A.; Underwood, A.; Williams, K. R.; Bassell, G. J.; Mihailescu, M. R. Fragile X mental retardation protein interactions with a G quadruplex structure in the 3′-untranslated region of NR2B mRNA. *Mol. BioSyst.* **2015**, *11* (12), 3222−3230.
(73) Poewe, W.; Seppi, K.; Tanner, C. M.; Halliday, G. M.; Brundin, P.; Volkmann, J.; Schrag, A.-E.; Lang, A. E. Parkinson disease. *Nat. Rev. Dis. Primers* **2017**, *3* (1), 17013.
(74) Sellier, C.; Buijsen, R. A. M.; He, F.; Natla, S.; Jung, L.; Tropel, P.; Gaucherot, A.; Jacobs, H.; Meziane, H.; Vincent, A.; Champy, M.-F.; Sorg, T.; Pavlovic, G.; Wattenhofer-Donze, M.; Birling, M.-C.; Oulad-Abdelghani, M.; Eberling, P.; Ruffenach, F.; Joint, M.; Anheim, M.; Martinez-Cerdeno, V.; Tassone, F.; Willemsen, R.; Hukema, R. K.; Viville, S.; Martinat, C.; Todd, P. K.; Charlet-Berguerand, N. Translation of Expanded CGG Repeats into FMRpolyG Is Pathogenic and May Contribute to Fragile X Tremor Ataxia Syndrome. *Neuron* **2017**, *93* (2), 331−347.
(75) Turcotte, M.-A.; Garant, J.-M.; Cossette-Roberge, H.; Perreault, J.-P. Guanine Nucleotide-Binding Protein-Like 1 (GNL1) binds RNA G-quadruplex structures in genes associated with Parkinson's disease. *RNA Biol.* **2021**, *18*, 1339.
(76) Selkoe, D. J. Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiol. Rev.* **2001**, *81* (2), 741−766.
(77) Yin, R.-H.; Yu, J.-T.; Tan, L. The Role of SORL1 in Alzheimer's Disease. *Mol. Neurobiol.* **2015**, *51* (3), 909−918.
(78) Weiss, P.; Murdoch, D. R. Clinical course and mortality risk of severe COVID-19. *Lancet* **2020**, *395* (10229), 1014−1015.
(79) Bohalova, N.; Cantara, A.; Bartas, M.; Kaura, P.; Stastny, J.; Pecinka, P.; Fojta, M.; Mergny, J.-L.; Brazda, V. Analyses of viral

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

genomes for G-quadruplex forming sequences reveal their correlation with the type of infection. *Biochimie* **2021**, *186*, 13−27.
(80) Lavezzo, E.; Berselli, M.; Frasson, I.; Perrone, R.; Palù, G.; Brazzale, A. R.; Richter, S. N.; Toppo, S. J. P. c. b. G-quadruplex forming sequences in the genome of all known human viruses: A comprehensive guide. *PLoS Comput. Biol.* **2018**, *14* (12), No. e1006675.
(81) Ruggiero, E.; Richter, S. N. G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res.* **2018**, *46* (7), 3270−3283.
(82) Saranathan, N.; Vivekanandan, P. G-Quadruplexes: More Than Just a Kink in Microbial Genomes. *Trends Microbiol.* **2019**, *27* (2), 148−163.
(83) Ruggiero, E.; Richter, S. N. Viral G-quadruplexes: New frontiers in virus pathogenesis and antiviral therapy. In *Quadruplex Nucleic Acids as Targets for Medicinal Chemistry*; Neidle, S., Ed.; Elsevier: 2020; Vol. *54*, pp 101−131.
(84) Abiri, A.; Lavigne, M.; Rezaei, M.; Nikzad, S.; Zare, P.; Mergny, J.-L.; Rahimi, H.-R. J. P. R. Unlocking G-Quadruplexes as Antiviral Targets. *Pharmacol. Rev.* **2021**, *73* (3), 897−923.
(85) Lan, K.; Luo, M.-H. Herpesviruses: epidemiology, pathogenesis, and interventions. *Virol. Sin.* **2017**, *32* (5), 347−348.
(86) McCORMICK, L.; Roller, R. J.; Roizman, B. Characterization of a herpes simplex virus sequence which binds a cellular protein as either a single-stranded or double-stranded DNA or RNA. *J. Virol.* **1992**, *66* (6), 3435−3447.
(87) Biswas, B.; Kandpal, M.; Jauhari, U. K.; Vivekanandan, P. Genome-wide analysis of G-quadruplexes in herpesvirus genomes. *BMC Genomics* **2016**, *17*, 949.
(88) Artusi, S.; Nadai, M.; Perrone, R.; Biasolo, M. A.; Palu, G.; Flamand, L.; Calistri, A.; Richter, S. N. The Herpes Simplex Virus-1 genome contains multiple clusters of repeated G-quadruplex: Implications for the antiviral activity of a G-quadruplex ligand. *Antiviral Res.* **2015**, *118*, 123−131.
(89) Biswas, B.; Kumari, P.; Vivekanandan, P. Pac1 Signals of Human Herpesviruses Contain a Highly Conserved G-Quadruplex Motif. *ACS Infect. Dis.* **2018**, *4* (5), 744−751.
(90) Frasson, I.; Nadai, M.; Richter, S. N. Conserved G-Quadruplexes Regulate the Immediate Early Promoters of Human Alphaherpesviruses. *Molecules* **2019**, *24* (13), 2375.
(91) Frasson, I.; Soldà, P.; Nadai, M.; Lago, S.; Richter, S. N. Parallel G-quadruplexes recruit the HSV-1 transcription factor ICP4 to promote viral transcription in herpes virus-infected human cells. *J. Biol. Chem.* **2021**, *4* (1), 1−13.
(92) Artusi, S.; Ruggiero, E.; Nadai, M.; Tosoni, B.; Perrone, R.; Ferino, A.; Zanin, I.; Xodo, L.; Flamand, L.; Richter, S. N. Antiviral Activity of the G-Quadruplex Ligand TMPyP4 against Herpes Simplex Virus-1. *Viruses* **2021**, *13* (2), 196.
(93) Artusi, S.; Perrone, R.; Lago, S.; Raffa, P.; Di Iorio, E.; Palu, G.; Richter, S. N. Visualization of DNA G-quadruplexes in herpes simplex virus 1-infected cells. *Nucleic Acids Res.* **2016**, *44* (21), 10343−10353.
(94) Tluckova, K.; Marusic, M.; Tothova, P.; Bauer, L.; Sket, P.; Plavec, J.; Viglasky, V. Human Papillomavirus G-Quadruplexes. *Biochemistry* **2013**, *52* (41), 7207−7216.
(95) Marusic, M.; Plavec, J. Towards Understanding of Polymorphism of the G-rich Region of Human Papillomavirus Type 52. *Molecules* **2019**, *24* (7), 1294.
(96) Thompson, A. B.; Flowers, L. C., Human papillomavirus (HPV). In *Sexually Transmitted Infections in Adolescence and Young Adulthood*; Springer: 2020; pp 279−297.
(97) Butovskaya, E.; Heddi, B.; Bakalar, B.; Richter, S. N.; Anh Tuan, P. Major G-Quadruplex Form of HIV-1 LTR Reveals a (3 + 1) Folding Topology Containing a Stem-Loop. *J. Am. Chem. Soc.* **2018**, *140* (42), 13654−13662.
(98) De Nicola, B.; Lech, C. J.; Heddi, B.; Regmi, S.; Frasson, I.; Perrone, R.; Richter, S. N.; Anh Tuan, P. Structure and possible function of a G-quadruplex in the long terminal repeat of the proviral HIV-1 genome. *Nucleic Acids Res.* **2016**, *44* (13), 6442−6451.
(99) Tosoni, E.; Frasson, I.; Scalabrin, M.; Perrone, R.; Butovskaya, E.; Nadai, M.; Palu, G.; Fabris, D.; Richter, S. N. Nucleolin stabilizes G-quadruplex structures folded by the LTR promoter and silences HIV-1 viral transcription. *Nucleic Acids Res.* **2015**, *43* (18), 8884−8897.
(100) Amrane, S.; Kerkour, A.; Bedrat, A.; Vialet, B.; Andreola, M.-L.; Mergny, J.-L. Topology of a DNA G-quadruplex structure formed in the HIV-1 promoter: a potential target for anti-HIV drug development. *J. Am. Chem. Soc.* **2014**, *136* (14), 5249−5252.
(101) Piekna-Przybylska, D.; Sullivan, M. A.; Sharma, G.; Bambara, R. A. U3 region in the HIV-1 genome adopts a G-quadruplex structure in its RNA and DNA sequence. *Biochemistry* **2014**, *53* (16), 2581−2593.
(102) Ruggiero, E.; Tassinari, M.; Perrone, R.; Nadai, M.; Richter, S. N. Stable and conserved G-quadruplexes in the long terminal repeat promoter of retroviruses. *ACS Infect. Dis.* **2019**, *5* (7), 1150−1159.
(103) Bartas, M.; Brazda, V.; Bohalova, N.; Cantara, A.; Volna, A.; Stachurova, T.; Malachova, K.; Jagelska, E. B.; Porubiakova, O.; Cerven, J.; Pecinka, P. In-Depth Bioinformatic Analyses of Nidovirales Including Human SARS-CoV-2, SARS-CoV, MERS-CoV Viruses Suggest Important Roles of Non-canonical Nucleic Acid Structures in Their Lifecycles. *Front. Microbiol.* **2020**, *11*, 1583.
(104) Cui, H.; Zhang, L. G-Quadruplexes Are Present in Human Coronaviruses Including SARS-CoV-2. *Front. Microbiol.* **2020**, *11*, 567317.
(105) Zhang, R.; Xiao, K.; Gu, Y.; Liu, H.; Sun, X. Whole Genome Identification of Potential G-Quadruplexes and Analysis of the G-Quadruplex Binding Domain for SARS-CoV-2. *Front. Genet.* **2020**, *11*, 587929.
(106) Belmonte-Reche, E.; Serrano-Chacón, I.; Gonzalez, C.; Gallo, J.; Bañobre-López, M. J. P. o. Potential G-quadruplexes and i-Motifs in the SARS-CoV-2. *PLoS One* **2021**, *16* (6), No. e0250654.
(107) Ji, D.; Juhas, M.; Tsang, C. M.; Kwok, C. K.; Li, Y.; Zhang, Y. Discovery of G-quadruplex-forming sequences in SARS-CoV-2. *Briefings Bioinf.* **2021**, *22* (2), 1150−1160.
(108) Zhao, C.; Qin, G.; Niu, J.; Wang, Z.; Wang, C.; Ren, J.; Qu, X. Targeting RNA G-Quadruplex in SARS-CoV-2: A Promising Therapeutic Target for COVID-19? *Angew. Chem., Int. Ed.* **2021**, *60* (1), 432−438.
(109) Brazda, V.; Porubiakova, O.; Cantara, A.; Bohalova, N.; Coufal, J.; Bartas, M.; Fojta, M.; Mergny, J.-L. G-quadruplexes in H1N1 influenza genomes. *BMC Genomics* **2021**, *22* (1), 77.
(110) Tomaszewska, M.; Szabat, M.; Zielińska, K.; Kierzek, R. Identification and Structural Aspects of G-Quadruplex-Forming Sequences from the Influenza A Virus Genome. *Int. J. Mol. Sci.* **2021**, *22* (11), 6031.
(111) Piekna-Przybylska, D.; Sharma, G.; Bambara, R. A. Mechanism of HIV-1 RNA dimerization in the central region of the genome and significance for viral evolution. *J. Biol. Chem.* **2013**, *288* (33), 24140−24150.
(112) Ravichandran, S.; Kim, Y.-E.; Bansal, V.; Ghosh, A.; Hur, J.; Subramani, V. K.; Pradhan, S.; Lee, M. K.; Kim, K. K.; Ahn, J.-H. Genome-Wide Analysis of Regulatory G-Quadruplexes Affecting Gene Expression in Human Cytomegalovirus. *PLoS Pathog.* **2018**, *14* (9), No. e1007334.
(113) Madireddy, A.; Purushothaman, P.; Loosbroock, C. P.; Robertson, E. S.; Schildkraut, C. L.; Verma, S. C. G-quadruplex-interacting compounds alter latent DNA replication and episomal persistence of KSHV. *Nucleic Acids Res.* **2016**, *44* (8), 3675−3694.
(114) Dabral, P.; Babu, J.; Zareie, A.; Verma, S. C. LANA and hnRNP A1 Regulate the Translation of LANA mRNA through G-Quadruplexes. *J. Virol.* **2020**, *94* (3), e01508-19.
(115) Gilbert-Girard, S.; Gravel, A.; Artusi, S.; Richter, S. N.; Wallaschek, N.; Kaufer, B. B.; Flamand, L. Stabilization of Telomere G-Quadruplexes Interferes with Human Herpesvirus 6A Chromosomal Integration. *J. Virol.* **2017**, *91* (14), No. e00402-17.
(116) Granzhan, A.; Martins, R. P.; Fåhraeus, R.; Blondel, M.; Teulade-Fichou, M. P. Quadruplex-interacting compounds for regulating the translation of the Epstein−Barr virus nuclear antigen

1 (EBNA1) mRNA: A new strategy to prevent and treat EBV-related cancers. In *Quadruplex Nucleic Acids As Targets For Medicinal Chemistry*; Stephen, N., Ed.; Elsevier: 2020; Vol. *54*, pp 243−286.
(117) Murat, P.; Zhong, J.; Lekieffre, L.; Cowieson, N. P.; Clancy, J. L.; Preiss, T.; Balasubramanian, S.; Khanna, R.; Tellam, J. G-quadruplexes regulate Epstein-Barr virus-encoded nuclear antigen 1 mRNA translation. *Nat. Chem. Biol.* **2014**, *10* (5), 358−U64.
(118) Norseen, J.; Johnson, F. B.; Lieberman, P. M. Role for G-Quadruplex RNA Binding by Epstein-Barr Virus Nuclear Antigen 1 in DNA Replication and Metaphase Chromosome Attachment. *J. Virol.* **2009**, *83* (20), 10336−10346.
(119) Lista, M. J.; Martins, R. P.; Billant, O.; Contesse, M.-A.; Findakly, S.; Pochard, P.; Daskalogianni, C.; Beauvineau, C.; Guetta, C.; Jamin, C.; Teulade-Fichou, M.-P.; Fahraeus, R.; Voisset, C.; Blondel, M. Nucleolin directly mediates Epstein-Barr virus immune evasion through binding to G-quadruplexes of EBNA1 mRNA. *Nat. Commun.* **2017**, *8* (1), 16043.
(120) Tellam, J. T.; Zhong, J.; Lekieffre, L.; Bhat, P.; Martinez, M.; Croft, N. P.; Kaplan, W.; Tellam, R. L.; Khanna, R. mRNA Structural Constraints on EBNA1 Synthesis Impact on In Vivo Antigen Presentation and Early Priming of CD8(+) T Cells. *PLoS Pathog.* **2014**, *10* (10), e1004423.
(121) Perrone, R.; Nadai, M.; Frasson, I.; Poe, J. A.; Butovskaya, E.; Smithgall, T. E.; Palumbo, M.; Palu, G.; Richter, S. N. A Dynamic G-Quadruplex Region Regulates the HIV-1 Long Terminal Repeat Promoter. *J. Med. Chem.* **2013**, *56* (16), 6521−6530.
(122) Duverger, A.; Wolschendorf, F.; Zhang, M.; Wagner, F.; Hatcher, B.; Jones, J.; Cron, R. Q.; van der Sluis, R. M.; Jeeninga, R. E.; Berkhout, B.; Kutsch, O. An AP-1 binding site in the enhancer/core element of the HIV-1 promoter controls the ability of HIV-1 to establish latent infection. *J. Virol.* **2013**, *87* (4), 2264−2277.
(123) Piekna-Przybylska, D.; Sharma, G.; Maggirwar, S. B.; Bambara, R. A. Deficiency in DNA damage response, a new characteristic of cells infected with latent HIV-1. *Cell Cycle* **2017**, *16* (10), 968−978.
(124) Piekna-Przybylska, D.; Bambara, R. A.; Maggirwar, S. B.; Dewhurst, S. G-quadruplex ligands targeting telomeres do not inhibit HIV promoter activity and cooperate with latency reversing agents in killing latently infected cells. *Cell Cycle* **2020**, *19* (18), 2298−2313.
(125) Piekna-Przybylska, D.; Maggirwar, S. B. CD4+ memory T cells infected with latent HIV-1 are susceptible to drugs targeting telomeres. *Cell Cycle* **2018**, *17* (17), 2187−2203.
(126) Bian, W.-X.; Xie, Y.; Wang, X.-N.; Xu, G.-H.; Fu, B.-S.; Li, S.; Long, G.; Zhou, X.; Zhang, X.-L. Binding of cellular nucleolin with the viral core RNA G-quadruplex structure suppresses HCV replication. *Nucleic Acids Res.* **2019**, *47* (1), 56−68.
(127) Shen, L.-W.; Qian, M.-Q.; Yu, K.; Narva, S.; Yu, F.; Wu, Y.-L.; Zhang, W. Inhibition of Influenza A virus propagation by benzoselenoxanthenes stabilizing TMPRSS2 Gene G-quadruplex and hence down-regulating TMPRSS2 expression. *Sci. Rep.* **2020**, *10* (1), 7635.
(128) Biswas, B.; Kandpal, M.; Vivekanandan, P. A G-quadruplex motif in an envelope gene promoter regulates transcription and virion secretion in HBV genotype B. *Nucleic Acids Res.* **2017**, *45* (19), 11268−11280.
(129) Meier-Stephenson, V.; Badmalia, M. D.; Mrozowich, T.; Lau, K. C.; Schultz, S. K.; Gemmill, D. L.; Osiowy, C.; van Marle, G.; Coffin, C. S.; Patel, T. R. Identification and characterization of a G-quadruplex structure in the pre-core promoter region of hepatitis B virus covalently closed circular DNA. *J. Biol. Chem.* **2021**, *296*, 100589−100589.
(130) Puig Lombardi, E.; Londoño-Vallejo, A.; Nicolas, A. J. M. Relationship between G-quadruplex sequence composition in viruses and their hosts. *Molecules* **2019**, *24* (10), 1942.
(131) Bohálová, N.; Cantara, A.; Bartas, M.; Kaura, P.; Šťastný, J.; Pečinka, P.; Fojta, M.; Brázda, V. Tracing dsDNA Virus−Host Coevolution through Correlation of Their G-Quadruplex-Forming Sequences. *Int. J. Mol. Sci.* **2021**, *22* (7), 3433.
(132) Sundquist, W. I.; Heaphy, S. Evidence for interstrand quadruplex formation in the dimerization of human immunodeficiency virus-1 genomic RNA. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90* (8), 3393−3397.
(133) Saranathan, N.; Biswas, B.; Patra, A.; Vivekanandan, P. G-quadruplexes may determine the landscape of recombination in HSV-1. *BMC Genomics* **2019**, *20* (1), 382.
(134) Tan, J.; Vonrhein, C.; Smart, O. S.; Bricogne, G.; Bollati, M.; Kusov, Y.; Hansen, G.; Mesters, J. R.; Schmidt, C. L.; Hilgenfeld, R. The SARS-Unique Domain (SUD) of SARS Coronavirus Contains Two Macrodomains That Bind G-Quadruplexes. *PLoS Pathog.* **2009**, *5* (5), No. e1000428.
(135) Kusov, Y.; Tan, J.; Alvarez, E.; Enjuanes, L.; Hilgenfeld, R. A G-quadruplex-binding macrodomain within the "SARS-unique domain" is essential for the activity of the SARS-coronavirus replication-transcription complex. *Virology* **2015**, *484*, 313−322.
(136) Bezzi, G.; Piga, E. J.; Binolfi, A.; Armas, P. CNBP Binds and Unfolds In Vitro G-Quadruplexes Formed in the SARS-CoV-2 Positive and Negative Genome Strands. *Int. J. Mol. Sci.* **2021**, *22* (5), 2614.
(137) Scalabrin, M.; Frasson, I.; Ruggiero, E.; Perrone, R.; Tosoni, E.; Lago, S.; Tassinari, M.; Palù, G.; Richter, S. N. The cellular protein hnRNP A2/B1 enhances HIV-1 transcription by unfolding LTR promoter G-quadruplexes. *Sci. Rep.* **2017**, *7* (1), 45244.
(138) Simone, R.; Balendra, R.; Moens, T. G.; Preza, E.; Wilson, K. M.; Heslegrave, A.; Woodling, N. S.; Niccoli, T.; Gilbert-Jaramillo, J.; Abdelkarim, S.; Clayton, E. L.; Clarke, M.; Konrad, M.-T.; Nicoll, A. J.; Mitchell, J. S.; Calvo, A.; Chio, A.; Houlden, H.; Polke, J. M.; Ismail, M. A.; Stephens, C. E.; Vo, T.; Farahat, A. A.; Wilson, W. D.; Boykin, D. W.; Zetterberg, H.; Partridge, L.; Wray, S.; Parkinson, G.; Neidle, S.; Patani, R.; Fratta, P.; Isaacs, A. M. G-quadruplex-binding small molecules ameliorate C9orf72 FTD/ALS pathology in vitro and in vivo. *EMBO Mol. Med.* **2018**, *10* (1), 22−31.
(139) Goldberg, D. C.; Fones, L.; Vivinetto, A. L.; Caufield, J. T.; Ratan, R. R.; Cave, J. W. Manipulating Adult Neural Stem and Progenitor Cells with G-Quadruplex Ligands. *ACS Chem. Neurosci.* **2020**, *11* (10), 1504−1518.
(140) Moruno-Manchon, J. F.; Lejault, P.; Wang, Y.; McCauley, B.; Honarpisheh, P.; Morales Scheihing, D. A.; Singh, S.; Dang, W.; Kim, N.; Urayama, A.; Zhu, L.; Monchaud, D.; McCullough, L. D.; Tsvetkov, A. S. Small-molecule G-quadruplex stabilizers reveal a novel pathway of autophagy regulation in neurons. *eLife* **2020**, *9*, No. e52283.
(141) Shioda, N.; Yabuki, Y.; Yamaguchi, K.; Onozato, M.; Li, Y.; Kurosawa, K.; Tanabe, H.; Okamoto, N.; Era, T.; Sugiyama, H.; Wada, T.; Fukunaga, K. Targeting G-quadruplex DNA as cognitive function therapy for ATR-X syndrome. *Nat. Med.* **2018**, *24* (6), 802−813.
(142) Rouleau, S.; Jodoin, R.; Garant, J.-M.; Perreault, J.-P. RNA G-Quadruplexes as Key Motifs of the Transcriptome. In *Catalytically Active Nucleic Acids*; Seitz, H., Stahl, F.; Walter, J. G., Eds.; Springer: 2020; Vol. *170*, pp 1−20.
(143) Cadoni, E.; De Paepe, L.; Manicardi, A.; Madder, A. Beyond small molecules: targeting G-quadruplex structures with oligonucleotides and their analogues. *Nucleic Acids Res.* **2021**, *49* (12), 6638−6659.
(144) Hagihara, M.; Yamauchi, L.; Seo, A.; Yoneda, K.; Senda, M.; Nakatani, K. Antisense-Induced Guanine Quadruplexes Inhibit Reverse Transcription by HIV-1 Reverse Transcriptase. *J. Am. Chem. Soc.* **2010**, *132* (32), 11171−11178.
(145) Sarkar, S.; Armitage, B. A. J. A. I. D. Targeting a Potential G-Quadruplex Forming Sequence Found in the West Nile Virus Genome by Complementary Gamma-Peptide Nucleic Acid Oligomers. *ACS Infect. Dis.* **2021**, *7* (6), 1445−1456.
(146) Donnelly, C. J.; Zhang, P.-W.; Pham, J. T.; Haeusler, A. R.; Mistry, N. A.; Vidensky, S.; Daley, E. L.; Poth, E. M.; Hoover, B.; Fines, D. M.; Maragakis, N.; Tienari, P. J.; Petrucelli, L.; Traynor, B. J.; Wang, J.; Rigo, F.; Bennett, C. F.; Blackshaw, S.; Sattler, R.; Rothstein, J. D. RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **2013**, *80* (2), 415−428.

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

(147) Lagier-Tourenne, C.; Baughn, M.; Rigo, F.; Sun, S.; Liu, P.; Li, H.-R.; Jiang, J.; Watt, A. T.; Chun, S.; Katz, M.; Qiu, J.; Sun, Y.; Ling, S.-C.; Zhu, Q.; Polymenidou, M.; Drenner, K.; Artates, J. W.; McAlonis-Downes, M.; Markmiller, S.; Hutt, K. R.; Pizzo, D. P.; Cady, J.; Harms, M. B.; Baloh, R. H.; Vandenberg, S. R.; Yeo, G. W.; Fu, X.-D.; Bennett, C. F.; Cleveland, D. W.; Ravits, J. Targeted degradation of sense and antisense *C9orf72* RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (47), E4530.

(148) Jiang, J.; Zhu, Q.; Gendron, T. F.; Saberi, S.; McAlonis-Downes, M.; Seelman, A.; Stauffer, J. E.; Jafar-nejad, P.; Drenner, K.; Schulte, D.; Chun, S.; Sun, S.; Ling, S.-C.; Myers, B.; Engelhardt, J.; Katz, M.; Baughn, M.; Platoshyn, O.; Marsala, M.; Watt, A.; Heyser, C. J.; Ard, M. C.; De Muynck, L.; Daughrity, L. M.; Swing, D. A.; Tessarollo, L.; Jung, C. J.; Delpoux, A.; Utzschneider, D. T.; Hedrick, S. M.; de Jong, P. J.; Edbauer, D.; Van Damme, P.; Petrucelli, L.; Shaw, C. E.; Bennett, C. F.; Da Cruz, S.; Ravits, J.; Rigo, F.; Cleveland, D. W.; Lagier-Tourenne, C. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in C9ORF72 Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron* **2016**, *90* (3), 535−550.

(149) Gendron, T. F.; Chew, J.; Stankowski, J. N.; Hayes, L. R.; Zhang, Y.-J.; Prudencio, M.; Carlomagno, Y.; Daughrity, L. M.; Jansen-West, K.; Perkerson, E. A.; O'Raw, A.; Cook, C.; Pregent, L.; Belzil, V.; van Blitterswijk, M.; Tabassian, L. J.; Lee, C. W.; Yue, M.; Tong, J.; Song, Y.; Castanedes-Casey, M.; Rousseau, L.; Phillips, V.; Dickson, D. W.; Rademakers, R.; Fryer, J. D.; Rush, B. K.; Pedraza, O.; Caputo, A. M.; Desaro, P.; Palmucci, C.; Robertson, A.; Heckman, M. G.; Diehl, N. N.; Wiggs, E.; Tierney, M.; Braun, L.; Farren, J.; Lacomis, D.; Ladha, S.; Fournier, C. N.; McCluskey, L. F.; Elman, L. B.; Toledo, J. B.; McBride, J. D.; Tiloca, C.; Morelli, C.; Poletti, B.; Solca, F.; Prelle, A.; Wuu, J.; Jockel-Balsarotti, J.; Rigo, F.; Ambrose, C.; Datta, A.; Yang, W.; Raitcheva, D.; Antognetti, G.; McCampbell, A.; Van Swieten, J. C.; Miller, B. L.; Boxer, A. L.; Brown, R. H.; Bowser, R.; Miller, T. M.; Trojanowski, J. Q.; Grossman, M.; Berry, J. D.; Hu, W. T.; Ratti, A.; Traynor, B. J.; Disney, M. D.; Benatar, M.; Silani, V.; Glass, J. D.; Floeter, M. K.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L. Poly(GP) proteins are a useful pharmacodynamic marker for *C9ORF72*-associated amyotrophic lateral sclerosis. *Sci. Transl. Med.* **2017**, *9* (383), No. eaai7866.

(150) Linden, R.; Martins, V. R.; Prado, M. A. M.; Cammarota, M.; Izquierdo, I.; Brentani, R. R. Physiology of the Prion Protein. *Physiol. Rev.* **2008**, *88* (2), 673−728.

(151) Mashima, T.; Nishikawa, F.; Kamatari, Y. O.; Fujiwara, H.; Saimura, M.; Nagata, T.; Kodaki, T.; Nishikawa, S.; Kuwata, K.; Katahira, M. Anti-prion activity of an RNA aptamer and its structural basis. *Nucleic Acids Res.* **2013**, *41* (2), 1355−1362.

(152) Shum, K. T.; Tanner, J. A. Differential Inhibitory Activities and Stabilisation of DNA Aptamers against the SARS Coronavirus Helicase. *ChemBioChem* **2008**, *9* (18), 3037−3045.

(153) Jang, K. J.; Lee, N.-R.; Yeo, W.-S.; Jeong, Y.-J.; Kim, D.-E. Isolation of inhibitory RNA aptamers against severe acute respiratory syndrome (SARS) coronavirus NTPase/Helicase. *Biochem. Biophys. Res. Commun.* **2008**, *366* (3), 738−744.

(154) Adli, M. The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* **2018**, *9* (1), 1911.

(155) Li, S.-Y.; Cheng, Q.-X.; Liu, J.-K.; Nie, X.-Q.; Zhao, G.-P.; Wang, J. CRISPR-Cas12a has both cis- and trans-cleavage activities on single-stranded DNA. *Cell Res.* **2018**, *28* (4), 491−493.

(156) Li, Y.; Li, T.; Liu, B.-F.; Hu, R.; Zhu, J.; He, T.; Zhou, X.; Li, C.; Yang, Y.; Liu, M. CRISPR-Cas12a trans-cleaves DNA G-quadruplexes. *Chem. Commun.* **2020**, *56* (83), 12526−12529.

(157) Balci, H.; Globyte, V.; Joo, C. Targeting G-quadruplex Forming Sequences with Cas9. *ACS Chem. Biol.* **2021**, *16* (4), 596−603.

(158) Wang, Z.; Cui, W. CRISPR-Cas system for biomedical diagnostic platforms. *VIEW* **2020**, *1* (3), 20200008.

(159) Polstein, L. R.; Gersbach, C. A. A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **2015**, *11* (3), 198−200.

(160) Wang, S.-R.; Wu, L.-Y.; Huang, H.-Y.; Xiong, W.; Liu, J.; Wei, L.; Yin, P.; Tian, T.; Zhou, X. Conditional control of RNA-guided nucleic acid cleavage and gene editing. *Nat. Commun.* **2020**, *11* (1), 91.

(161) Baran, N.; Pucshansky, L.; Marco, Y.; Benjamin, S.; Manor, H. The SV40 large T-antigen helicase can unwind four stranded DNA structures linked by G-quartets. *Nucleic Acids Res.* **1997**, *25* (2), 297−303.

(162) Mendoza, O.; Bourdoncle, A.; Boule, J.-B.; Brosh, R. M., Jr.; Mergny, J.-L. G-quadruplexes and helicases. *Nucleic Acids Res.* **2016**, *44* (5), 1989−2006.

(163) Prentice, E.; McAuliffe, J.; Lu, X. T.; Subbarao, K.; Denison, M. R. Identification and characterization of severe acute respiratory syndrome coronavirus replicase proteins. *J. Virol.* **2004**, *78* (18), 9977−9986.

(164) Briguglio, I.; Piras, S.; Corona, P.; Carta, A. Inhibition of RNA Helicases of ssRNA(+) Virus Belonging to Flaviviridae, Coronaviridae and Picornaviridae Families. *Int. J. Med. Chem.* **2011**, *2011*, 213135−213135.

(165) Tanner, J. A.; Zheng, B. J.; Zhou, J.; Watt, R. M.; Jiang, J. Q.; Wong, K. L.; Lin, Y. P.; Lu, L. Y.; He, M. L.; Kung, H. F.; Kesel, A. J.; Huang, J. D. The adamantane-derived bananins are potent inhibitors of the helicase activities and replication of SARS coronavirus. *Chem. Biol.* **2005**, *12* (3), 303−311.

(166) Cho, J.-B.; Lee, J.-M.; Ahn, H.-C.; Jeong, Y.-J. Identification of a Novel Small Molecule Inhibitor Against SARS Coronavirus Helicase. *J. Microbiol. Biotechnol.* **2015**, *25* (12), 2007−2010.

(167) Liu, H.; Lu, Y.-N.; Paul, T.; Periz, G.; Banco, M. T.; Ferré-D'Amaré, A. R.; Rothstein, J. D.; Hayes, L. R.; Myong, S.; Wang, J. A Helicase Unwinds Hexanucleotide Repeat RNA G-Quadruplexes and Facilitates Repeat-Associated Non-AUG Translation. *J. Am. Chem. Soc.* **2021**, *143* (19), 7368−7379.

(168) Cheng, W.; Wang, S.; Zhang, Z.; Morgens, D. W.; Hayes, L. R.; Lee, S.; Portz, B.; Xie, Y.; Nguyen, B. V.; Haney, M. S.; Yan, S.; Dong, D.; Coyne, A. N.; Yang, J.; Xian, F.; Cleveland, D. W.; Qiu, Z.; Rothstein, J. D.; Shorter, J.; Gao, F.-B.; Bassik, M. C.; Sun, S. CRISPR-Cas9 Screens Identify the RNA Helicase DDX3X as a Repressor of C9ORF72 (GGGGCC)n Repeat-Associated Non-AUG Translation. *Neuron* **2019**, *104* (5), 885−898.

(169) Lee, D. S. M.; Ghanem, L. R.; Barash, Y. Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. *Nat. Commun.* **2020**, *11* (1), 527.

(170) Collie, G. W.; Parkinson, G. N.; Neidle, S.; Rosu, F.; De Pauw, E.; Gabelica, V. Electrospray mass spectrometry of telomeric RNA (TERRA) reveals the formation of stable multimeric G-quadruplex structures. *J. Am. Chem. Soc.* **2010**, *132* (27), 9328−9334.

(171) Huang, H.-Y.; Wang, S.-R.; Wu, L.-Y.; Zhang, Y.; Song, Y.; Wei, L.; Tian, T.; Zhou, X. J. C. C. Biochemical Insights into the Role of Guanosine Oxidation on RNA G-Quadruplex. *CCS Chemistry* **2020**, *2* (6), 605−612.

(172) Shrestha, P.; Jonchhe, S.; Emura, T.; Hidaka, K.; Endo, M.; Sugiyama, H.; Mao, H. Confined space facilitates G-quadruplex formation. *Nat. Nanotechnol.* **2017**, *12* (6), 582−588.

(173) Herdy, B.; Mayer, C.; Varshney, D.; Marsico, G.; Murat, P.; Taylor, C.; D'Santos, C.; Tannahill, D.; Balasubramanian, S. Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. *Nucleic Acids Res.* **2018**, *46* (21), 11592−11604.

(174) Hacht, A. v.; Seifert, O.; Menger, M.; Schütze, T.; Arora, A.; Konthur, Z.; Neubauer, P.; Wagner, A.; Weise, C.; Kurreck, J. Identification and characterization of RNA guanine-quadruplex binding proteins. *Nucleic Acids Res.* **2014**, *42* (10), 6630−6644.

(175) Su, H.; Xu, J.; Chen, Y.; Wang, Q.; Lu, Z.; Chen, Y.; Chen, K.; Han, S.; Fang, Z.; Wang, P.; Yuan, B.-F.; Zhou, X. Photoactive G-Quadruplex Ligand Identifies Multiple G-Quadruplex-Related Pro-

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161

teins with Extensive Sequence Tolerance in the Cellular Environment.
J. Am. Chem. Soc. **2021**, *143* (4), 1917−1923.

https://doi.org/10.1021/jacsau.1c00451
JACS Au 2021, 1, 2146−2161